

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
18 December 2003 (18.12.2003)

PCT

(10) International Publication Number  
**WO 03/104391 A2**

- (51) International Patent Classification<sup>7</sup>: C12N
- (21) International Application Number: PCT/US02/36122
- (22) International Filing Date:  
25 November 2002 (25.11.2002)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:  
60/333,777 29 November 2001 (29.11.2001) US  
60/426,742 18 November 2002 (18.11.2002) US  
60/427,367 18 November 2002 (18.11.2002) US
- (71) Applicant (for all designated States except US): **WYETH HOLDINGS CORPORATION** [US/US]; a corporation of the State of Maine, Five Giralda Farms, Madison, NJ 07940-0874 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): **MURPHY, Ellen** [US/US]; 185 Beach Street, City Island, NY 10464 (US). **PROJAN, Steven, Jay** [US/US]; 104 Gedney Street, Nyack, NY 10960 (US).
- (74) Agents: **CALNAN, William, H.** et al.; Wyeth, Patent Law Department, Five Giralda Farms, Madison, NJ 07940-874 (US).
- (81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.
- (84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- Published:**  
— without international search report and to be republished upon receipt of that report
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.



WO 03/104391 A2

(54) Title: ANTIBACTERIAL TARGETS IN ALLOIOCOCCUS OTITIDIS

(57) Abstract: The present invention relates to the identification of polynucleotide sequences encoding polypeptides of *Alloiococcus otitidis* that are essential for the growth and survival of the bacteria. In particular, the invention relates to polypeptides encoded by the *Alloiococcus otitidis* open reading frames (ORFs), and to their use in pharmaceutical compositions, therapeutics, diagnostics and the like. The present invention also relates to methods for identifying pharmaceutical compounds that inhibit the activity of the polypeptides that are essential for the growth of *Alloiococcus otitidis*, to pharmaceutical compositions containing these compounds and to their use in treatment and amelioration of diseases caused by *Alloiococcus otitidis*.

## ANTIBACTERIAL TARGETS IN ALLOIOCOCCUS OTITIDIS

### FIELD OF THE INVENTION

5 The present invention relates to the genomic sequence of *Alloiococcus otitidis* and polynucleotide sequences encoding polypeptides of the Gram-positive bacterium, *Alloiococcus otitidis*. The invention also relates to polynucleotides and polynucleotides encoding polypeptides, preferably antigenic polypeptides, encoded by the *Alloiococcus otitidis* open reading frames and the uses thereof.

### BACKGROUND OF THE INVENTION

10 Since the discovery of penicillin, the use of antibiotics to treat the ravages of bacterial infections has saved millions of lives. With the advent of these "miracle drugs," for a time it was popularly believed that humanity might, once and for all, be saved from the scourge of bacterial infections. In fact, during the 1980s and early 15 1990s, many large pharmaceutical companies cut back or eliminated antibiotics research and development. They believed that infectious disease caused by bacteria finally had been conquered and that markets for new drugs were limited. Unfortunately, this belief was overly optimistic. The tide is beginning to turn in favor of the bacteria, as reports of drug resistant bacteria become more frequent. The United States Centers for Disease Control and Prevention announced that one of the most powerful known antibiotics, vancomycin, was unable to treat an infection of the common bacterial pathogen, *Staphylococcus aureus*. This organism, commonly found in our environment, is responsible for many nosocomial infections. The import 25 of this announcement becomes clear when one considers that vancomycin was used for years to treat infections caused by *Staphylococcus* species as well as other stubborn strains of bacteria. In short, bacteria are becoming resistant to our most powerful antibiotics. If this trend continues, it is conceivable that we will return to a time when what are presently considered minor bacterial infections are fatal diseases. 30

Over-prescription and improper prescription habits by some physicians have caused an indiscriminate increase in the availability of antibiotics to the public. The patients are also partly responsible, since they will often improperly use the drug,

thereby generating yet another population of bacteria that is resistant, in whole or in part, to traditional antibiotics.

The bacterial pathogens that have haunted humanity remain, in spite of the development of modern scientific practices to deal with the diseases that they cause.

5 Drug resistant bacteria are now an increasing threat to the health of humanity. A new generation of antibiotics is needed to once again deal with the pending health threats that bacteria present.

As more and more bacterial strains become resistant to the panel of available antibiotics, new antibiotics are required to treat infections. In the past, practitioners of  
10 pharmacology relied upon traditional methods of drug discovery to generate novel, safe and efficacious compounds for the treatment of disease. Traditional drug discovery methods involve blindly testing potential drug candidate- molecules, often selected at random, in the hope that one might prove to be an effective treatment for some disease. The process is painstaking and laborious, with no guarantee of  
15 success.

Newly emerging practices in drug discovery utilize a number of biochemical techniques to provide for directed approaches to creating new drugs, rather than discovering them at random. For example, gene sequences and proteins encoded thereby that are required for the proliferation of a cell or microorganism make  
20 excellent targets since exposure of bacteria to compounds active against these targets would result in the inactivation of the cell or microorganism. Once a target is identified, biochemical analysis of that target can be used to discover or to design molecules that interact with and alter the functions of the target. Use of physical and computational techniques to analyze structural and biochemical properties of targets  
25 in order to derive compounds that interact with such targets is called rational drug design and offers great potential. Thus, emerging drug discovery practices use molecular modeling techniques, combinatorial chemistry approaches, and other means to produce and screen and/or design large numbers of candidate compounds.

Nevertheless, while this approach to drug discovery is clearly the way of the  
30 future, problems remain. For example, the initial step of identifying molecular targets for investigation can be an extremely time consuming task. It may also be difficult to design molecules that interact with the target by using computer modeling techniques. Furthermore, in cases where the function of the target is not known or is

poorly understood, it may be difficult to design assays to detect molecules that interact with and alter the functions of the target. To improve the rate of novel drug discovery and development, methods of identifying important molecular targets in pathogenic cells or microorganisms and methods for identifying molecules that interact with and alter the functions of such molecular targets are urgently required.

The present invention is directed to identifying important molecular targets in a recently identified bacteria, *Alloiococcus otitidis*, which has been implicated in otitis media with effusion (OME). Otitis media, an inflammatory disease of the middle ear, is the most frequent cause of visits to pediatricians' offices in the United States (Schappert, 1991). Approximately 80% of all children experience at least one episode of otitis media by the age of three (Klein, 1994). There are three main types of otitis media: Acute otitis media (AOM), otorrhea, and otitis media with effusion (OME). *Alloiococcus otitidis* has only been associated with otitis media with effusion (OME), but this may be due to the difficulty of its detection by standard bacterial culturing methods. Its detection in the effusions is likely due to the fact that the effusions are normally sterile and few or no competing bacterial species are isolated from them. Without the interference of faster growing nasopharyngeal species, the culture plates can be incubated for the longer duration needed to detect *Alloiococcus otitidis* colonies.

Three other bacterial species are commonly isolated from middle ear effusions. These are nontypeable *Haemophilus influenzae*, *Moraxella catarrhalis*, and *Streptococcus pneumoniae*. One or more of these species have been found in one study to be associated with about 77% of all cases of OME using a PCR detection method (Post, 2000). This study did not include assaying for *Alloiococcus otitidis*, so a portion of the unaccounted cases may be due to this organism.

The bacterium *Alloiococcus otitidis* was first isolated from the middle ear fluids of 10 children in the Buffalo, NY area with persistent OME and characterized as a large catalase negative, Gram-positive cocci that tend to occur in clumps, often in tetrads. It is slow growing and requires 2 to 5 days at 37°C before colonies can be seen on sheep blood agar plates. The bacterium was named *Alloiococcus otitis* by Aguirre and Collins (1992), who showed that it was different from other known Gram-positive species based on its 16S rRNA sequence. The bacterium's name has been



changed from *Alloiococcus otitis* to *Alloiococcus otitidis*. (Hendolin, et al., (1999), and Hendolin et al., (2000)).

Several studies of the epidemiology *Alloiococcus otitidis* indicate it is associated with otitis media with effusion. These are summarized in Table 1. These studies have been done using both culture and PCR techniques. The number of cases detected by culture, as might be expected from the fastidious growth requirements of the bacterium, was less than the number detected by PCR. Assuming that the bacterium is detected more accurately by the PCR method, the bacterium is detected in between 10 and 50% of patients with OME. This frequency suggests that this organism represents a significant public health problem. Consequently, there is a need for identifying gene targets in *Alloiococcus otitidis* for the development of anti-infectives. There is also a need for compositions for diagnosing *Alloiococcus otitidis* infection.

**TABLE 1: SUMMARY OF STUDIES INDICATING AN ASSOCIATION OF ALLOIOCOCCUS OTITIDIS WITH OTITIS MEDIA WITH EFFUSION (OME).**

% detected	N <sup>a</sup>	Method	Reference
8	200	Culture	Faden & Dryja, J. Clin. Microbiol. 27:2488 (1989)
3	100	Culture	Sih et al., ICAAC (1992)
20	25	PCR	Hendolin et al., J. Clin. Microbiol. 35:2854 (1997)
50	12	PCR	Beswick, et al., Lancet 345:386 (1999)
42	67	PCR	Hendolin, et al., Pediatr. Infect. Dis. J. 18:860 (1999)
10	49	PCR	Hendolin et al., J. Clin. Microbiol. 38:125 (2000)

<sup>a</sup> Number of persons in study.

### SUMMARY OF INVENTION

The present invention broadly relates to *Alloiococcus otitidis* genomic sequence. Particularly, the invention relates to newly identified polynucleotide open reading frames (ORFs) comprised within the genomic nucleotide sequence of *Alloiococcus otitidis*, and to polypeptides encoded by the ORFs. More particularly, the ORFs encode polypeptides that are essential for the growth and survivability of *Alloiococcus otitidis*.

Thus, in certain aspects, the invention relates to *Alloiococcus otitidis* ORFs that encode *Alloiococcus otitidis* polypeptides that function as enzymes in various biosynthetic pathways in the bacterium. In one embodiment, the invention relates to a purified or isolated *Alloiococcus otitidis* nucleic acid sequence comprising a nucleotide sequence selected from one of odd numbered sequences set forth in Seq. ID Nos: 1 to Seq. ID Nos: 105, wherein expression of said nucleic acid is essential for the proliferation of a cell. In a preferred embodiment the ORF selected from one of the odd numbered sequence listings set forth in Seq. ID Nos: 1 to Seq. ID Nos: 105 encodes an essential gene. The essential gene and the polypeptide encoded by them include ACPS (holo-(acyl carrier protein) synthase), murF (UDP-N-acetylmuramoylalanyl-D-glutamyl-2,6-diamino pimelate-D-alanyl-D-alanyl ligase) murA-2 (UDP-N-acetylglucosamine 1-carboxyvinyltransferase), RpoE (DNA-directed RNA polymerase, delta subunit), rpoA (DNA-directed RNA polymerase alpha subunit), rpoC (RNA polymerase beta' subunit), rpoB (DNA-dependent RNA polymerase subunit beta), dnaB/C (DNA polymerase III delta prime subunit), gyrA (DNA gyrase A subunit), gyrB (DNA gyrase B subunit), dnaN (DNA polymerase III beta chain, folC-2 (folyl-polyglutamate synthetase), murE (UDP-N-acetylmuramoyl-L-alanyl-D-glutamyl-L-lysine Ligase), srtA (sortase), folC-1 (folyl-polyglutamate synthetase), folB (dihydroneopterin aldolase), folK (7,8-dihydro-6-hydroxymethylpterin-pyrophosphokinase), mvaS (hydroxymethylglutaryl-CoA synthase), mvaA (3-hydroxy-3-methylglutaryl-coenzyme a reductase), murB (UDP-N-acetylglucosaminyl-3-enolpyruvate reductase), mvaK2 (phosphomevalonate kinase), mvaD (mevalonate diphosphate decarboxylase), mvaK1 (mevalonate kinase), coaA (pantothenate kinase), nadE (NAD<sup>+</sup> synthase), murl, Glutamate racemase), folP (Dihydropteroate synthase), folA (dihydrofolate reductase), grlB (topoisomerase IV B

subunit), *grlA* (topoisomerase IV A subunit), *rpoD* (transcription initiation factor sigma), *dnaG* (DNA primase), *era* (GTP-binding protein), *norA* (drug-export protein), *polC* (DNA polymerase III, alpha subunit), *obg* (GTP-binding protein), *yphC* (similar to *Escherichia coli* GTP-binding protein Era), *dnaE* (DNA polymerase III, alpha subunit), *coaBC* (phosphopantothienoylcysteine synthetase/decarboxylase), *holA* (DNA polymerase III delta subunit), *coaD* (phosphopantetheine adenylyltransferase) *ftsZ* (Cell division protein *ftsZ*), *ftsA* (Cell division protein *ftsA*), *murG* (phospho-N-acetylmuramoyl-pentapeptide-transferase), *murD* (UDP-N-acetylmuramoylalanine D-glutamate ligase), *nadD* (nicotinic acid mononucleotide adenylyltransferase), *coaE* (dephospho-CoA kinase), *murC* (UDP-N-acetyl muramate-alanine ligase), *fmhB* *FemX* (factor essential for methicillin resistance), *pcrA* (ATP-dependent DNA helicase), *murA-1* (UDP-N-acetylglucosamine 1-carboxyvinyltransferase), *holB* (DNA polymerase III delta' subunit) and *dnaX* (DNA polymerase III -gamma and tau subunits).

15 In another embodiment, the invention relates to purified or isolated nucleic acid of *Alloiococcus otitidis* comprising a fragment of one of odd numbered sequences set forth in Seq. ID Nos: 1 to Seq. ID Nos: 105, wherein said fragment is selected from the group consisting of fragments comprising at least 10, at least 20, at least 25, at least 30, at least 50 and more than 50 consecutive nucleotides of one of one of odd numbered sequences set forth in Seq. ID Nos: 1 to Seq. ID Nos: 105.

20 In yet another embodiment, the invention relates to a purified or isolated antisense nucleic acid comprising a nucleotide sequence complementary to at least a portion of an intragenic sequence, intergenic sequence, sequences spanning at least a portion of two or more genes, 5' noncoding region, or 3' noneoding region within an operon comprising a proliferation-required gene of *Alloiococcus otitidis* whose activity or expression is inhibited by an antisense nucleic acid and selected from one of odd numbered sequences set forth in Seq. ID Nos: 1 to Seq. ID Nos: 105.

25 In a nother embodiment, the invention relates to a purified or isolated nucleic acid comprising a nucleotide sequence having at least 70% identity to a nucleotide sequence selected from one of odd numbered sequences set forth in Seq. ID Nos: 1 to Seq. ID Nos: 105, fragments comprising at least 25 consecutive nucleotides selected from one of odd numbered sequences set forth in Seq. ID Nos: 1 to Seq. ID Nos: 105, the nucleotide sequences complementary to one of odd numbered

sequences set forth in Seq. ID Nos: 1 to Seq. ID Nos: 105, and the sequences complementary to fragments comprising at least 25 consecutive nucleotides of one of odd numbered sequences set forth in Seq. ID Nos: 1 to Seq. ID Nos: 105.

5 In another embodiment, the invention relates to a vector comprising a promoter operably linked to a nucleic acid encoding a polypeptide whose expression is inhibited by an antisense nucleic acid comprising a nucleotide sequence of any one of odd numbered sequences set forth in Seq. ID Nos: 1 to Seq. ID Nos: 105.

10 In another embodiment, the invention relates to purified or isolated polypeptide of *Alloicoccus otitidis* comprising a polypeptide whose expression is inhibited by an antisense nucleic acid comprising a nucleotide sequence of one of odd numbered sequences set forth in Seq. ID Nos: 1 to Seq. ID Nos: 105, or a fragment selected from the group consisting of fragments comprising at least 5, at least 10, at least 20, at least 30, at least 40, at least 50, at least 60 or more than 60 consecutive amino acids of one of the said polypeptides.

15 In yet another embodiment, the invention relates to purified or isolated *Alloicoccus otitidis* polypeptide comprising a amino acid sequence having at least 25% amino acid identity to a polypeptide whose expression is inhibited by a nucleic acid comprising a nucleotide sequence selected from one of odd numbered sequences set forth in Seq. ID Nos: 1 to Seq. ID Nos: 105, or at least 25% amino acid identity to a fragment comprising at least 10, at least 20, at least 30, at least 40, at least 50, at least 60 or more than 60 consecutive amino acids of a polypeptide whose expression is inhibited by a nucleic acid comprising a nucleotide sequence selected from the group consisting of one of odd numbered sequences set forth in Seq. ID Nos: 1 to Seq. ID Nos: 105.

25 In one embodiment, the invention relates to a purified or isolated *Alloicoccus otitidis* polypeptide comprising selected from one of the even numbered sequences set forth in Seq. ID Nos: 2 to Seq. ID Nos: 106, wherein the polypeptide is essential for the proliferation of a cell..

30 In yet another embodiment, the invention relates to a method of producing an *Alloicoccus otitidis* polypeptide comprising introducing into a cell a vector comprising a promoter operably linked to a nucleic acid comprising a nucleotide sequence encoding a polypeptide whose expression is essential for the proliferation and viability of *Alloicoccus otitidis*, and which is inhibited by an antisense nucleic

acid, and which is selected from one of odd numbered sequences set forth in Seq. ID Nos: 1 to Seq. ID Nos: 105.

5 In yet another embodiment, the invention relates to a method of inhibiting the proliferation of *Alloicoccus otitidis* in an individual comprising inhibiting the activity or reducing the amount of a gene product whose expression is inhibited by an antisense nucleic acid comprising a nucleotide sequence selected from one of odd numbered sequences set forth in Seq. ID Nos: 1 to Seq. ID Nos: 105 or inhibiting the activity or reducing the amount of a nucleic acid encoding said gene product.

10 In a preferred embodiment, the invention relates to method for identifying a compound which influences the activity of an *Alloicoccus otitidis* gene product, which is required for proliferation, said gene product comprising a gene product whose expression is inhibited by an antisense nucleic acid comprising a nucleotide sequence selected from one of odd numbered sequences set forth in Seq. ID Nos: 1 to Seq. ID Nos: 105, said method comprising: (a) contacting said gene product with a candidate compound; and (b) determining whether said compound influences the activity of said gene product.

20 In a preferred embodiment, the invention relates to method for identifying a compound or an antisense nucleic acid having the ability to reduce activity or level of a *Alloicoccus otitidis* gene product, which is required for proliferation, said gene product comprising a gene product whose activity or expression is inhibited by an antisense nucleic acid comprising a nucleotide sequence selected from one of odd numbered sequences set forth in Seq. ID Nos: 1 to Seq. ID Nos: 105, said method comprising the steps of: (a) contacting a target gene or RNA encoding said gene product with a candidate compound or antisense nucleic acid; and (b) measuring the activity of said target.

30 In yet another preferred embodiment, the invention relates to method for inhibiting cellular proliferation of *Alloicoccus otitidis* comprising introducing an effective amount of a compound with activity against a gene whose activity or expression is essential for cellular proliferation, and which is inhibited by an antisense nucleic acid comprising a nucleotide sequence selected from one of odd numbered sequences set forth in Seq. ID Nos: 1 to Seq. ID Nos: 105, or a compound with activity against the product of said gene into a population of *Alloicoccus otitidis* cells expressing said gene.

In a preferred embodiment, the invention relates to a composition comprising an effective concentration of an antisense nucleic acid comprising a nucleotide sequence selected from one of odd numbered sequences set forth in Seq. ID Nos: 1 to Seq. ID Nos: 105, or a proliferation-inhibiting portion thereof in a pharmaceutically acceptable carrier.

In a preferred embodiment, the invention relates to method for identifying a compound having the ability to inhibit proliferation of *Alloiococcus otitidis* cell comprising: (a) identifying a homologue of a gene or gene product whose activity or level is inhibited by a nucleic acid comprising a nucleotide sequence selected from one of odd numbered sequences set forth in Seq. ID Nos: 1 to Seq. ID Nos: 105, in a test cell, wherein said test cell is not *Alloiococcus otitidis*; (b) identifying an inhibitory nucleic acid sequence which inhibits the activity of said homologue in said test cell; (c) contacting said test cell with a sublethal level of said inhibitory nucleic acid, thus sensitizing said cell; (d) contacting the sensitized cell of step (c) with a compound; and (e) determining the degree to which said compound inhibits proliferation of said sensitized cell relative to a cell which does not contain said inhibitory nucleic acid.

In a preferred embodiment, the invention relates to a method for identifying a compound having activity against a biological pathway required for proliferation comprising: (a) sensitizing a cell by providing a sublethal level of an antisense nucleic acid complementary to a nucleic acid encoding a gene product required for proliferation, wherein the activity or expression of said gene product is inhibited by an antisense nucleic acid comprising a nucleotide sequence selected from one of odd numbered sequences set forth in Seq. ID Nos: 1 to Seq. ID Nos: 105, in said cell to reduce the activity or amount of said gene product; (b) contacting the sensitized cell with a compound; and (c) determining the degree to which said compound inhibits the growth of said sensitized cell relative to a cell which does not contain said antisense nucleic acid.

In a preferred embodiment, the invention relates to a method for identifying a compound having the ability to inhibit one of the *Alloiococcus otitidis* polypeptides encoded by a polynucleotide selected from one of odd numbered sequences set forth in Seq. ID Nos: 1 to Seq. ID Nos: 105, and which is essential for cellular proliferation comprising: (a) contacting a cell which expresses the polypeptide with the compound;

and (b) determining whether said compound reduces proliferation of said contacted cell by acting on said gene product.

In a preferred embodiment, the invention relates to a method for identifying a compound having the ability to inhibit one of the purified and isolated *Alloicoccus* *otitidis* polypeptides selected from one of the even numbered sequences set forth in Seq. ID No.: 2 to Seq. ID No.: 106, and which is essential for cellular proliferation comprising: (a) contacting the purified and isolated polypeptide with the compound *in vitro* in the presence or absence of a substrate, which is essential for the activity of the polypeptide; and (b) determining the effect of the compound on the polypeptide by measuring the effect of the polypeptide on the substrate.

In a preferred embodiment, the invention relates to a compound which interacts with an *Alloicoccus otitidis* polypeptide selected from one of the even numbered sequences set forth in Seq. ID No.: 2 to Seq. ID No.: 106 and inhibits its activity.

In a preferred embodiment, the invention relates to a method for manufacturing an antimicrobial compound comprising the steps of screening one or more candidate compounds to identify a compound that reduces the activity or level of an *Alloicoccus otitidis* polypeptide selected from one of the even numbered sequences set forth in Seq. ID No.: 2 to Seq. ID No.: 106, said polypeptide comprising a gene product whose activity or expression is inhibited by an antisense nucleic acid comprising a nucleotide sequence selected from one of the odd numbered sequences set forth in Seq. ID No.: 1 to Seq. ID No. 105; and manufacturing the compound so identified.

In a preferred embodiment, the invention relates to a compound which inhibits proliferation of *Alloicoccus otitidis* by interacting with a gene encoding a polypeptide that is required for proliferation or with a polypeptide required for proliferation, wherein said polypeptide is selected from the group consisting of a gene product having at least 70% nucleotide sequence identity from one of the odd numbered sequences set forth in Seq. ID No.: 1 to Seq. ID No. 105, polypeptide encoded by a nucleic acid having at least 70% nucleotide sequence identity to a nucleic acid encoding a polypeptide whose expression is inhibited by an antisense nucleic acid comprising a nucleotide sequence selected from one of the odd numbered sequences set forth in Seq. ID No.: 1 to Seq. ID No. 105, a polypeptide having at

least 25% amino acid identity to a gene product whose expression is inhibited by an antisense nucleic acid comprising a nucleotide sequence selected one of the odd numbered sequences set forth in Seq. ID No.: 1 to Seq. ID No. 105, a polypeptide encoded by a nucleic acid comprising a nucleotide sequence which hybridizes to a nucleic acid selected from one of the odd numbered sequences set forth in Seq. ID No.: 1 to Seq. ID No. 105 under stringent conditions, a gene product encoded by a nucleic acid comprising a nucleotide sequence which hybridizes to a nucleic acid selected from one of the odd numbered sequences set forth in Seq. ID No.: 1 to Seq. ID No. 105 under moderate conditions, and a gene product whose activity may be complemented by the gene product whose activity is inhibited by a nucleic acid selected from one of the odd numbered sequences set forth in Seq. ID No.: 1 to Seq. ID No. 105.

#### DETAILED DESCRIPTION OF THE INVENTION

15

##### **A. Definitions:**

By "biological pathway" is meant any discrete cell function or process that is carried out by a gene product or a subset of gene products. Biological pathways include anabolic, catabolic, enzymatic, biochemical and metabolic pathways as well as pathways involved in the production of cellular structures such as cell walls. Biological pathways that are usually required for proliferation of cells or microorganisms include, but are not limited to, cell division, DNA synthesis and replication, RNA synthesis (transcription), protein synthesis (translation), protein processing, protein transport, fatty acid biosynthesis, electron transport chains, cell wall synthesis, cell membrane production, synthesis and maintenance, and the like.

By "inhibit activity of a gene or gene product" is meant having the ability to interfere with the function of a gene or gene product in such a way as to decrease expression of the gene, in such a way as to reduce the level or activity of a product of the gene or in such a way as to inhibit the interaction of the gene or gene product with other biological molecules required for its activity.

30

Agents which inhibit the activity of a gene include agents that inhibit transcription of the gene, agents that inhibit processing of the transcript of the gene, agents that reduce the stability of the transcript of the gene, and agents that inhibit translation of the mRNA transcribed from the gene. In microorganisms, agents which



inhibit the activity of a gene can act to decrease expression of the operon in which the gene resides or alter the folding or processing of operon RNA so as to reduce the level or activity of the gene product. The gene product can be a non- translated RNA such as ribosomal RNA, a translated RNA (mRNA) or the protein product resulting from translation of the gene mRNA. Of particular utility to the present invention are antisense RNAs that have activities against the operons or genes to which they specifically hybridize.

By "activity against a gene product" is meant having the ability to inhibit the function or to reduce the level or activity of the gene product in a cell. This includes, but is not limited to, inhibiting the enzymatic activity of the gene product or the ability of the gene product to interact with other biological molecules required for its activity, including inhibiting the gene product's assembly into a multimeric structure.

By "activity against a protein" is meant having the ability to inhibit the function or to reduce the level or activity of the protein in a cell. This includes, but is not limited to, inhibiting the enzymatic activity of the protein or the ability of the protein to interact with other biological molecules required for its activity, including inhibiting the protein's assembly into a multimeric structure.

By "activity against a nucleic acid" is meant having the ability to inhibit the function or to reduce the level or activity of the nucleic acid in a cell. This includes, but is not limited to, inhibiting the ability of the nucleic acid interact with other biological molecules required for its activity, including inhibiting the nucleic acid's assembly into a multimeric structure.

By "activity against a gene" is meant having the ability to inhibit the function or expression of the gene in a cell. This includes, but is not limited to, inhibiting the ability of the gene to interact with other biological molecules required for its activity.

By "activity against an operon" is meant having the ability to inhibit the function or reduce the level of one or more products of the operon in a cell. This includes, but is not limited to, inhibiting the enzymatic activity of one or more products of the operon or the ability of one or more products of the operon to interact with other biological molecules required for its activity.

By "antibiotic" is meant an agent which inhibits the proliferation of a cell or microorganism.

By "homologous coding nucleic acid" is meant a nucleic acid homologous to a nucleic acid encoding a gene product whose activity or level is inhibited by a nucleic acid selected from the group consisting of Seq ID Nos.: 1 to Seq. ID Nos.: 105 or a portion thereof. In some embodiments, the homologous coding nucleic acid may

5 have at least 97%, at least 95%, at least 90%, at least 85%, at least 80%, or at least 70% nucleotide sequence identity to a nucleotide sequence selected from the group consisting of Seq ID Nos.: 1 to Seq. ID Nos.: 105 and fragments comprising at least 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, 150, 200, 300, 400, or 500 consecutive nucleotides thereof. In other embodiments the homologous coding nucleic acids may

10 have at least 97%, at least 95%, at least 90%, at least 85%, at least 80%, or at least 70% nucleotide sequence identity to a nucleotide sequence selected from the group consisting of the nucleotide sequences complementary to one of Seq ID Nos.: 1 to Seq. ID Nos.: 105 and fragments comprising at least 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, 150, 200, 300, 400, or 500 consecutive nucleotides thereof. Identity may

15 be measured using BLASTN version 2.0 with the default parameters or tBLASTX with the default parameters. (Altschul, S.F. et al. Gapped BLAST and PSI-BLAST: A New Generation of Protein Database Search Programs, *Nucleic Acid Res.* 25: 3389-3402 (1997)) Alternatively a "homologous coding nucleic acid" could be identified by membership of the gene of interest to a functional orthologue cluster. All other

20 members of that orthologue cluster would be considered homologues. Such a library of functional orthologue clusters can be found at <http://www.ncbi.nlm.nih.gov/COG>. A gene can be classified into a cluster of orthologous groups or COG by using the COGNITOR program available at the above web site, or by direct BLASTP comparison of the gene of interest to the members of the COGs and analysis of

25 these results as described by Tatusov, R.L., Galperin, M.Y., Natale, D. A. and Koonin, E.V. (2000) The COG database: a tool for genome- scale analysis of protein functions and evolution. *Nucleic Acids Research* v. 28 n. 1, pp33-36.

The term "homologous coding nucleic acid" also includes nucleic acids comprising nucleotide sequences which encode polypeptides having at least 99%,

30 95%, at least 90%, at least 85%, at least 80%, at least 70%, at least 60%, at least 50%, at least 40% or at least 25% amino acid identity or similarity to a polypeptide comprising the amino acid sequence of one of Seq ID Nos.: 1 to Seq. ID Nos.: 105 or to a polypeptide whose expression is inhibited by a nucleic acid comprising a

nucleotide sequence of one of Seq ID Nos.: 1 to Seq. ID Nos.: 105 or fragments comprising at least 5, 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, or 150 consecutive amino acids thereof as determined using the FASTA version 3.0t78 algorithm with the default parameters. Alternatively, protein identity or similarity may be identified using BLASTP with the default parameters, BLASTX with the default parameters, TBLASTN with the default parameters, or tBLASTX with the default parameters. (Altschul, S.F. et al. Gapped BLAST and PSI-BLAST: A New Generation of Protein Database Search Programs, *Nucleic Acid Res.* 25: 3389-3402 (1997)).

The term "homologous coding nucleic acid" also includes coding nucleic acids which hybridize under stringent conditions to a nucleic acid selected from the group consisting of the nucleotide sequences complementary to one of Seq ID Nos.: 1 to Seq. ID Nos.: 105 and coding nucleic acids comprising nucleotide sequences which hybridize under stringent conditions to a fragment comprising at least 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, 150, 200, 300, 400, or 500 consecutive nucleotides of the sequences complementary to one of Seq ID Nos.: 1 to Seq. ID Nos.: 105.

As used herein, "stringent conditions" means hybridization to filter-bound nucleic acid in 6xSSC at about 45°C followed by one or more washes in 0.1xSSC/0.2% SDS at about 68°C. Other exemplary stringent conditions may refer, e.g., to washing in 6xSSC/0.05% sodium pyrophosphate at 37°C, 48°C, 55°C, and 60°C as appropriate for the 5 particular probe being used.

The term "homologous coding nucleic acid" also includes coding nucleic acids comprising nucleotide sequences which hybridize under moderate conditions to a nucleotide sequence selected from the group consisting of the sequences complementary to one of Seq ID Nos.: 1 to Seq. ID Nos.: 105 and coding nucleic acids comprising nucleotide sequences which hybridize under moderate conditions to a fragment comprising at least 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, 150, 200, 300, 400, or 500 consecutive nucleotides of the sequences complementary to one of Seq ID Nos.: 1 to Seq. ID Nos.: 105. As used herein, "moderate conditions" means hybridization to filter-bound DNA in 6x sodium chloride/sodium citrate (SSC) at about 45°C followed by one or more washes in 0.2xSSC/0.1% SDS at about 42-65°C.

The term "homologous coding nucleic acids" also includes nucleic acids comprising nucleotide sequences which encode a gene product whose activity may

be complemented by a gene encoding a gene product whose activity is inhibited by a nucleic acid comprising a nucleotide sequence selected from the group consisting of Seq ID Nos.: 1 to Seq. ID Nos.: 105. In some embodiments, the homologous coding nucleic acids may encode a gene product whose activity is complemented by the gene product encoded by a nucleic acid comprising a nucleotide sequence selected from the group consisting Seq ID Nos.: 1 to Seq. ID Nos.: 105. In other embodiments, the homologous coding nucleic acids may comprise a nucleotide sequence encodes a gene product whose activity is complemented by one of the polypeptides of Seq ID Nos.: 1 to Seq. ID Nos.: 105 .

The term "homologous antisense nucleic acid" includes nucleic acids comprising a nucleotide sequence having at least 97%, at least 95%, at least 90%, at least 85%, at least 80%, or at least 70% nucleotide sequence identity to a nucleotide sequence selected from the group consisting of one of the sequences of Seq ID Nos.: 1 to Seq. ID Nos.: 105 and fragments comprising at least 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, 150, 200, 300, 400, or 500 consecutive nucleotides thereof. Homologous antisense nucleic acids may also comprising nucleotide sequences which have at least 97%, at least 95%, at least 90%, at least 85%, at least 80%, or at least 70% nucleotide sequence identity to a nucleotide sequence selected from the group consisting of the sequences complementary to one of sequences of Seq ID Nos.: 1 to Seq. ID Nos.: 105 and fragments comprising at least 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, 150, 200, 300, 400, or 500 consecutive nucleotides thereof.

Nucleic acid identity may be determined as described above.

The term "homologous antisense nucleic acid" also includes antisense nucleic acids comprising nucleotide sequences which hybridize under stringent conditions to a nucleotide sequence complementary to one of Seq ID Nos.: 1 to Seq. ID Nos.: 105 and antisense nucleic acids comprising nucleotide sequences which hybridize under stringent conditions to a fragment comprising at least 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, 150, 200, 300, 400, or 500 consecutive nucleotides of the sequence complementary to one Seq ID Nos.: 1 to Seq. ID Nos.: 105. Homologous antisense nucleic acids also include antisense nucleic acids comprising nucleotide sequences which hybridize under stringent conditions to a nucleotide sequence selected from the group consisting of Seq ID Nos.: 1 to Seq. ID Nos.: 105, and antisense nucleic acids comprising nucleotide sequences which hybridize under stringent conditions to

a fragment comprising at least 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, 150, 200, 300, 400, or 500 consecutive nucleotides of one of Seq ID Nos.: 1 to Seq. ID Nos.: 105.

The term "homologous antisense nucleic acid" also includes antisense  
5 nucleic acids comprising nucleotide sequences which hybridize under moderate conditions to a nucleotide sequence complementary to one of Seq ID Nos.: 1 to Seq. ID Nos.: 105 and antisense nucleic acids comprising nucleotide sequences which hybridize under moderate conditions to a fragment comprising at least 10, 15, 20, 25,  
10 30, 35, 40, 50, 75, 100, 150, 200, 300, 400, or 500 consecutive nucleotides of the sequence complementary to one of Seq ID Nos.: 1 to Seq. ID Nos.: 105.

Homologous antisense nucleic acids also include antisense nucleic acids comprising nucleotide sequences which hybridize under moderate conditions to a nucleotide sequence selected from the group consisting of Seq ID Nos.: 1 to Seq. ID Nos.: 105 and antisense nucleic acids which comprising nucleotide sequences hybridize under  
15 moderate conditions to a fragment comprising at least 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, 150, 200, 300, 400, or 500 consecutive nucleotides of one of Seq ID Nos.: 1 to Seq. ID Nos.: 105.

By "homologous polypeptide" is meant a polypeptide homologous to a polypeptide whose activity or level is inhibited by a nucleic acid comprising a  
20 nucleotide sequence selected from the group consisting of Seq ID Nos.: 1 to Seq. ID Nos.: 105 by a homologous antisense nucleic acid. The term "homologous polypeptide" includes polypeptides having at least 99%, 95%, at least 90%, at least 85%, at least 80%, at least 70%, at least 60%, at least 50%, at least 40% or at least 25% amino acid identity or similarity to a polypeptide whose activity or level is  
25 inhibited by a nucleic acid selected from the group consisting of Seq ID Nos.: 1 to Seq. ID Nos.: 105 or by a homologous antisense nucleic acid, or polypeptides having at least 99%, 95%, at least 90%, at least 85%, at least 80%, at least 70%, at least 60%, at least 50%, at least 40% or at least 25% amino acid identity or similarity to a polypeptide to a fragment comprising at least 5, 10, 15, 20, 25, 30, 35, 40, 50, 75,  
30 100, or 150 consecutive amino acids of a polypeptide whose activity or level is inhibited by a nucleic acid selected from the group consisting of Seq ID Nos.: 1 to Seq. ID Nos.: 105 or by a homologous antisense nucleic acid. Identity or similarity may be determined using the FASTA version 3. Ot78 algorithm with the default

parameters. Alternatively, protein identity or similarity may be identified using BLASTP with the default parameters, BLASTX with the default parameters, or TBLASTN with the default parameters. (Altschul, S.F. et al. Gapped BLAST and PSI-BLAST: A New Generation of Protein Database Search Programs, Nucleic Acid Res. 25: 3389-3402 (1997).

The term homologous polypeptide also includes polypeptides having at least 99%, 95%, at least 90%, at least 85%, at least 80%, at least 70%, at least 60%, at least 50%, at least 40% or at least 25% amino acid identity or similarity to a polypeptide selected from the group consisting of Seq ID Nos.: 2 to Seq. ID Nos.: 106 and polypeptides having at least 99%, 95%, at least 90%, at least 85%, at least 80%, at least 70%, at least 60%, at least 50%, at least 40% or at least 25% amino acid identity or similarity to a fragment comprising at least 5, 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, or 150 consecutive amino acids of a polypeptide selected from the group consisting of Seq ID Nos.: 2 to Seq. ID Nos.: 106.

The invention also includes polynucleotides, preferably DNA molecules, that hybridize to one of the nucleic acids of Seq ID Nos.: 2 to Seq. ID Nos.: 106 or the complements of any of the preceding nucleic acids. Such hybridization may be under stringent or moderate conditions as defined above or under other conditions which permit specific hybridization. The nucleic acid molecules of the invention that hybridize to these DNA sequences include oligodeoxynucleotides ("oligos") which hybridize to the target gene under highly stringent or stringent conditions. In general, for oligos between 14 and 70 nucleotides in length the melting temperature ( $T_m$ ) is calculated using the formula:

$$T_m (^\circ\text{C}) = 81.5 + 16.6(\log[\text{monovalent cations (molar)}]) + 0.41 (\% \text{ G+C}) - (500/N)$$

where N is the length of the probe. If the hybridization is carried out in a solution containing formamide, the melting temperature may be calculated using the equation:

$$T_m (^\circ\text{C}) = 81.5 + 16.6(\log[\text{monovalent cations (molar)}]) + 0.41 (\% \text{ G+C}) - (0.61) (\% \text{ formamide}) - (500/N)$$

where N is the length of the probe. In general, hybridization is carried out at about 20-25 degrees below  $T_m$  (for DNA-DNA hybrids) or about 10-15 degrees below  $T_m$  (for RNA-DNA hybrids).

Other hybridization conditions are apparent to those of skill in the art (see, for example, Ausubel, F.M. et al., eds., 1989, Current Protocols in Molecular Biology,

Vol. 1, Green Publishing Associates, Inc. and John Wiley & Sons, Inc., New York, at pp. 6.3.1-6.3.6 and 2.10.3.

By "identifying a compound" is meant to screen one or more compounds in a collection of compounds such as a combinatorial chemical library or other library of chemical compounds or to characterize a single compound by testing the compound in a given assay and determining whether it exhibits the desired activity.

By "inducer" is meant an agent or solution which, when placed in contact with a cell or microorganism, increases transcription, or inhibitor and/or promoter clearance/fidelity, from a desired promoter.

As used herein, "nucleic acid" means DNA, RNA, or modified nucleic acids. Thus, the terminology "the nucleic acid of SEQ ID NO: V or "the nucleic acid comprising the nucleotide sequence" includes both the DNA sequence of SEQ ID NO: X and an RNA sequence in which the thymidines in the DNA sequence have been substituted with uridines in the RNA sequence and in which the deoxyribose backbone of the DNA sequence has been substituted with a ribose backbone in the RNA sequence. Modified nucleic acids are nucleic acids having nucleotides or structures which do not occur in nature, such as nucleic acids in which the internucleotide phosphate residues with methylphosphonates, phosphorothioates, phosphoramidates, and phosphate esters. Nonphosphate internucleotide analogs such as siloxane bridges, carbonate bridges, thioester bridges, as well as many others known in the art may also be used in modified nucleic acids. Modified nucleic acids may also comprise, (x-anomeric nucleotide units and modified nucleotides such as 1,2-dideoxy-d-ribofuranose, 1,2-dideoxy-1-phenylribofuranose, and N<sup>4</sup>, N<sup>4</sup>-ethano-5-methyl-cytosine are contemplated for use in the present invention. Modified nucleic acids may also be peptide nucleic acids in which the entire deoxyribose-phosphate backbone has been exchanged with a chemically completely different, but structurally homologous, polyamide (peptide) backbone containing 2-aminoethyl glycogen units.

As used herein, "sub-lethal" means a concentration of an agent below the concentration required to inhibit all cell growth.

A proliferation-required gene or gene family is one where, in the absence or substantial reduction of a gene transcript and/or gene product, growth or viability of the cell or microorganism is reduced or eliminated. Thus, as used herein, the

terminology "proliferation- required" or "required for proliferation" encompasses instances where the absence or substantial reduction of a gene transcript and/or gene product completely eliminates cell growth as well as instances where the absence of a gene transcript and/or gene product merely reduces cell growth. These proliferation-required genes can be used as potential targets for the generation of new antimicrobial agents. To achieve that goal, the present invention also encompasses assays for analyzing proliferation- required genes and for identifying compounds which interact with the gene and/or gene products of the proliferation-required genes. In addition, the present invention contemplates the expression of genes and the purification of the proteins encoded by the nucleic acid sequences identified as required proliferation genes and reported herein. The purified proteins can be used to generate reagents and screen small molecule libraries or other candidate compound libraries for compounds that can be further developed to yield novel antimicrobial compounds.

The invention described herein addresses the need for identifying *Alloiococcus otitidis* proliferation-required gene or gene family that may be used to identify compounds, which are effective in preventing or treating most or all of the disease caused by *Alloiococcus otitidis*. The invention further addresses the need for methods of diagnosing *Alloiococcus otitidis* infection using the genes and the polypeptides identified herein. The inventors have identified novel *Alloiococcus otitidis* open reading frames (Ors), which encode proteins/polypeptides that are essential for the growth and proliferation of the bacteria. More particularly, the newly identified Ors encode polypeptides that are essential for proliferation of *Alloiococcus otitidis*, and thus serve as potential targets for antimicrobial compounds. Thus, in certain embodiments, the invention comprises *Alloiococcus otitidis* Ors encoding polypeptides that are essential for cellular proliferation, transcription gene products of *Alloiococcus otitidis* Ors, including, but not limited to mRNA, antisense RNA, antisense oligonucleotides, and ribozyme molecules, which can be used to inhibit or control growth of the microorganism. The invention relates also to methods of detecting *Alloiococcus otitidis* nucleic acids or polypeptides and kits for diagnosing *Alloiococcus otitidis* infection. The invention also relates to pharmaceutical compositions, in particular antimicrobial compounds in pharmaceutical compositions, for the prevention and/or treatment of bacterial infection, in particular infection



caused by or exacerbated by *Alloiococcus otitidis*.

**B. ALLOIOCOCCUS OTTIDIS ORF POLYNUCLEOTIDES ENCODING POLYPEPTIDES  
ESSENTIAL FOR PROLIFERATION**

5

Isolated and purified *Alloiococcus otitidis* ORF polynucleotides of the present invention are contemplated for use in the production of *Alloiococcus otitidis* polypeptides. More specifically, in certain embodiments, the ORFs encode *Alloiococcus otitidis* polypeptides that are essential for cell proliferation. Thus, in one aspect, the present invention provides isolated and purified polynucleotides (ORFs) that encode *Alloiococcus otitidis* essential for cell proliferation. In particular embodiments, a polynucleotide of the present invention is a DNA molecule, wherein the DNA may be genomic DNA, plasmid DNA or cDNA. In a preferred embodiment, a polynucleotide of the present invention is a recombinant polynucleotide, which encodes an *Alloiococcus otitidis* polypeptide comprising an amino acid sequence that has at least 25% identity to an amino acid sequence of one of even numbered sequences set forth in SEQ ID NO: 2 through SEQ ID NO: 106 or a fragment thereof. In another embodiment, an isolated and purified ORF polynucleotide comprises a nucleotide sequence that has at least 70% identity to one of the ORF polynucleotide nucleotide sequences set forth in SEQ ID NO: 1 through SEQ ID NO: 105, a degenerate variant thereof, or a complement thereof. In yet another embodiment, an ORF polynucleotide of one of SEQ ID NO: 1 through SEQ ID NO: 105 is comprised in a plasmid vector and expressed in a host cell. In a preferred embodiment, the host cell is a prokaryotic host cell.

25

As used herein, the term "polynucleotide" means a sequence of nucleotides connected by phosphodiester linkages. Polynucleotides are presented herein in the direction from the 5' to the 3' direction. A polynucleotide of the present invention can comprise from about 10 to about several hundred thousand base pairs. Preferably, a polynucleotide comprises from about 10 to about 3,000 base pairs. Preferred lengths of particular polynucleotide are set forth hereinafter.

30

A polynucleotide of the present invention can be a deoxyribonucleic acid (DNA) molecule, a ribonucleic acid (RNA) molecule, or analogs of the DNA or RNA generated using nucleotide analogs. The nucleic acid molecule can be single-stranded or double-stranded, but preferably is double-stranded DNA. Where a

polynucleotide is a DNA molecule, that molecule can be a gene, a cDNA molecule or a genomic DNA molecule. Nucleotide bases are indicated herein by a single letter code: adenine (A), guanine (G), thymine (T) and cytosine (C).

"Isolated" means altered "by the hand of man" from the natural state. An "isolated" composition or substance is one that has been changed or removed from its original environment, or both. For example, a polynucleotide or a polypeptide naturally present in a living animal is not "isolated," but the same polynucleotide or polypeptide separated from the coexisting materials of its natural state is "isolated," as the term is employed herein.

Preferably, an "isolated" polynucleotide is free of sequences which naturally flank the nucleic acid (*i.e.*, sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated *Alloicoccus otitidis* nucleic acid molecule can contain less than about 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0.5 kb or 0.1 kb of nucleotide sequences which naturally flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived. However, the *Alloicoccus otitidis* nucleic acid molecule can also be fused to heterologous protein encoding or regulatory sequences and still be considered isolated.

ORF polynucleotides of the present invention may also be obtained using standard cloning and screening techniques from a cDNA library derived from mRNA. Polynucleotides of the invention can also be obtained from natural sources such as genomic DNA libraries (*e.g.*, an *Alloicoccus otitidis* library) or can be synthesized using well-known and commercially available techniques. As contemplated in the present invention, ORF polynucleotides are obtained using *Alloicoccus otitidis* chromosomal DNA as the template.

The invention further encompasses nucleic acid molecules that differ from the nucleotide sequences set forth in the odd numbered sequences listed in ID NO: 1 through SEQ ID NO: 105 (and fragments thereof) due to degeneracy of the genetic code, and thus encode the same *Alloicoccus otitidis* polypeptides as those encoded by the amino acid sequences shown in even numbered sequences set forth in SEQ ID NO:2 through SEQ ID NO: 106

Orthologs and allelic variants of the *Alloicoccus otitidis* polynucleotides are readily identified using methods well known in the art. An allelic variant or an

orthologue of the polynucleotides comprises a nucleotide sequence that is typically at least about 70-75%, more typically at least about 80-85%, and most typically at least about 90-95% or more homologous to the nucleotide sequence shown in one of the odd numbered sequences set forth in SEQ ID NO:1 through SEQ ID NO: 105, or a  
5 fragment of these nucleotide sequences. Such nucleic acid molecules are readily identified as being able to hybridize, preferably under stringent conditions, to the nucleotide sequence shown in one of the odd numbered sequences set forth in SEQ ID NO:1 through SEQ ID NO: 105, or a fragment of these nucleotide sequences.

Moreover, the polynucleotides of the invention can comprise only a fragment  
10 of the coding region of an *Alloicoccus otitidis* polynucleotide or gene, such as a fragment of one of the odd numbered sequences set forth in SEQ ID NO:1 through SEQ ID NO: 105.

When the ORF polynucleotides of the invention are used for the recombinant production of *Alloicoccus otitidis* polypeptides of the present invention, the  
15 polynucleotide may include the coding sequence for the mature polypeptide, by itself, or the coding sequence for the mature polypeptide in reading frame with other coding sequences, such as those encoding a leader or secretory sequence, a pre-, or pro- or prepro- protein sequence, or other fusion peptide portions. For example, a marker sequence which facilitates purification of the fused polypeptide can be linked to the  
20 coding sequence (see Gentz *et al.*, 1989, incorporated herein by reference). Thus, contemplated in the present invention is the preparation of polynucleotides encoding fusion polypeptides permitting His-tag purification of expression products. The polynucleotide may also contain non-coding 5' and 3' sequences, such as transcribed, non-translated sequences, splicing and polyadenylation signals.

Thus, a polynucleotide encoding a polypeptide of the present invention,  
25 including homologs and orthologs from species other than *Alloicoccus otitidis*, may be obtained by a process which comprises the steps of screening an appropriate library under stringent hybridization conditions with a labeled probe having the sequence of one of the odd numbered sequences set forth in SEQ ID NO:1 through  
30 SEQ ID NO: 105 or a fragment thereof; and isolating full-length cDNA and genomic clones containing the polynucleotide sequence. Such hybridization techniques are well known to the skilled artisan. The skilled artisan will appreciate that, in many cases, an isolated cDNA sequence will be incomplete, in that the region coding for

the polypeptide is cut short at the 5' end of the cDNA. This is a consequence of reverse transcriptase, an enzyme with inherently low "processivity" (a measure of the ability of the enzyme to remain attached to the template during the polymerization reaction), failing to complete a DNA copy of the mRNA template during the first-strand cDNA synthesis.

Thus, in certain embodiments, the polynucleotide sequence information provided by the present invention allows for the preparation of relatively short DNA (or RNA) oligonucleotide sequences having the ability to specifically hybridize to gene sequences of the selected polynucleotides disclosed herein. The term "oligonucleotide" as used herein is defined as a molecule comprised of two or more deoxyribonucleotides or ribonucleotides, usually more than three (3), and typically more than ten (10) and up to one hundred (100) or more (although preferably between twenty and thirty). The exact size will depend on many factors, which in turn depends on the ultimate function or use of the oligonucleotide. Thus, in particular embodiments of the invention, nucleic acid probes of an appropriate length are prepared based on a consideration of a selected nucleotide sequence, e.g., a sequence such as that shown in one of the odd numbered sequences set forth in SEQ ID NO:1 through SEQ ID NO: 105. The ability of such nucleic acid probes to specifically hybridize to a polynucleotide encoding an *Alloicoccus otitidis* polypeptide lends them particular utility in a variety of embodiments. Most importantly, the probes can be used in a variety of assays for detecting the presence of complementary sequences in a given sample.

In certain embodiments, it is advantageous to use oligonucleotide primers. These primers are generated in any manner, including chemical synthesis, DNA replication, reverse transcription, or a combination thereof. The sequence of such primers is designed using a polynucleotide of the present invention for use in detecting, amplifying or mutating a defined segment of an ORF polynucleotide that encodes an *Alloicoccus otitidis* polypeptide from prokaryotic cells using polymerase chain reaction (PCR) technology.

In certain embodiments, it is advantageous to employ a polynucleotide of the present invention in combination with an appropriate label for detecting hybrid formation. A wide variety of appropriate labels are known in the art, including

radioactive, enzymatic or other ligands, such as avidin/biotin, which are capable of giving a detectable signal.

Polynucleotides which are identical or sufficiently identical to a nucleotide sequence contained in one of the odd numbered sequences set forth in SEQ ID NO:1 through SEQ ID NO: 105, or a fragment thereof, may be used as hybridization probes for cDNA and genomic DNA or as primers for a nucleic acid amplification (PCR) reaction, to isolate full-length cDNAs and genomic clones encoding polypeptides of the present invention and to isolate cDNA and genomic clones of other genes (including genes encoding homologs and orthologs from species other than *Alloiococcus otitidis*) that have a high sequence similarity to polynucleotide sequences set forth in one of the odd numbered sequences set forth in SEQ ID NO:1 through SEQ ID NO:105, or a fragment thereof. Typically these nucleotide sequences are from at least 70% identical to at least about 95% identical to that of the reference polynucleotide sequence. The probes or primers will generally comprise at least 15 nucleotides, preferably, at least 30 nucleotides and may have at least 50 nucleotides. Particularly preferred probes will have between 30 and 50 nucleotides.

There are several methods available and well known to those skilled in the art to obtain full-length cDNAs, or extend short cDNAs, for example those based on the method of Rapid Amplification of cDNA ends (RACE) (*see, Frohman et al., 1988*). Recent modifications of the technique, exemplified by the Marathon™ technology [Promega, Madison, WI], for example, have significantly simplified the search for longer cDNAs. In the Marathon™ technology, cDNAs have been prepared from mRNA extracted from a chosen tissue and an "adaptor" sequence ligated onto each end. Nucleic acid amplification (PCR) is then carried out to amplify the "missing" 5' end of the cDNA using a combination of gene specific and adaptor specific oligonucleotide primers. The PCR reaction is then repeated using "nested" primers, that is, primers designed to anneal within the amplified product (typically an adaptor specific primer that anneals further 3' in the adaptor sequence and a gene specific primer that anneals further 5' in the known gene sequence). The products of this reaction are then analyzed by DNA sequencing and a full-length cDNA constructed either by joining the product directly to the existing cDNA to give a complete

sequence, or carrying out a separate full-length PCR using the new sequence information for the design of the 5' primer.

To provide certain of the advantages in accordance with the present invention, a preferred nucleic acid sequence employed for hybridization studies or assays includes probe molecules that are complementary to at least a 10 to about 70 nucleotides long stretch of a polynucleotide that encodes an *Alloiococcus otitidis* polypeptide, such as that shown in one of the even numbered sequences set forth in SEQ ID NO: 2 through SEQ ID NO: 106. A size of at least 10 nucleotides in length helps to ensure that the fragment will be of sufficient length to form a duplex molecule that is both stable and selective. Molecules having complementary sequences over stretches greater than 10 bases in length are generally preferred in order to increase stability and selectivity of the hybrid, and thereby improve the quality and degree of specific hybrid molecules obtained. It is generally preferable to design nucleic acid molecules with gene-complementary stretches of 25 to 40 nucleotides, 55 to 70 nucleotides, or even longer where desired. For example, such fragments are readily prepared by directly synthesizing the fragment by chemical means, by application of nucleic acid reproduction technology, such as the PCR technology (U.S. Patent 4,683,202, incorporated herein by reference), or by excising selected DNA fragments from recombinant plasmids containing appropriate inserts and suitable restriction enzyme sites.

In another aspect, the present invention contemplates an isolated and purified polynucleotide comprising a nucleotide sequence that is identical or complementary to a segment of at least 10 contiguous bases of one of the odd numbered sequences set forth in SEQ ID NO: 1 through SEQ ID NO: 105, wherein the polynucleotide hybridizes to a polynucleotide that encodes an *Alloiococcus otitidis* polypeptide. Preferably, the isolated and purified polynucleotide comprises a base sequence that is identical or complementary to a segment of at least 25 to 70 contiguous bases of one of the odd numbered sequences set forth in SEQ ID NO: 1 through SEQ ID NO: 105. For example, the polynucleotide of the invention can comprise a segment of bases identical or complementary to from 40 to 55 contiguous bases of the disclosed nucleotide sequences.

Accordingly, a polynucleotide probe molecule of the invention can be used for its ability to selectively form duplex molecules with complementary stretches of the

gene. Depending on the application envisioned, varying conditions of hybridization are employed to achieve varying degrees of selectivity of the probe toward the target sequence. For applications requiring a high degree of selectivity, relatively stringent conditions are employed to form the hybrids. Of course, for some applications, for example, where one desires to prepare mutants employing a mutant primer strand hybridized to an underlying template or where one seeks to isolate an *Alloiococcus otitidis* homologous polypeptide coding sequence from other cells, functional equivalents, or the like, less stringent hybridization conditions are typically needed to allow formation of the heteroduplex (see Table 2). Cross-hybridizing species are thereby readily identified as positively hybridizing signals with respect to control hybridizations. Thus, hybridization conditions are readily manipulated, and thus will generally be a method of choice depending on the desired results.

Of course, for some applications, for example, where one desires to prepare mutants employing a mutant primer strand hybridized to an underlying template or where one seeks to isolate a homologous polypeptide coding sequence from other cells, functional equivalents, or the like, less stringent hybridization conditions are typically needed to allow formation of the heteroduplex. Cross-hybridizing species are thereby readily identified as positively hybridizing signals with respect to control hybridizations. In any case, it is generally appreciated that conditions can be rendered more stringent by the addition of increasing amounts of formamide, which serves to destabilize the hybrid duplex in the same manner as increased temperature. Thus, hybridization conditions are readily manipulated, and thus are generally a method of choice depending on the desired results.

The present invention also includes polynucleotides capable of hybridizing under reduced stringency conditions, more preferably stringent conditions, and most preferably highly stringent conditions, to polynucleotides described herein. Examples of stringency conditions are shown in the table below: highly stringent conditions are those that are at least as stringent as, for example, conditions A-F; stringent conditions are at least as stringent as, for example, conditions G-L; and reduced stringency conditions are at least as stringent as, for example, conditions M-R.

**TABLE 2**  
**STRINGENCY CONDITIONS**

Stringency Condition	Polynucleotide Hybrid	Hybrid Length (bp) <sup>1</sup>	Hybridization Temperature and Buffer <sup>H</sup>	Wash Temperature and Buffer <sup>H</sup>
A	DNA:DNA	> 50	65°C; 1xSSC -or- 42°C; 1xSSC, 50% formamide	65°C; 0.3xSSC
B	DNA:DNA	< 50	T <sub>B</sub> ; 1xSSC	T <sub>B</sub> ; 1xSSC
C	DNA:RNA	> 50	67°C; 1xSSC -or- 45°C; 1xSSC, 50% formamide	67°C; 0.3xSSC
D	DNA:RNA	< 50	T <sub>D</sub> ; 1xSSC	T <sub>D</sub> ; 1xSSC
E	RNA:RNA	> 50	70°C; 1xSSC -or- 50°C; 1xSSC, 50% formamide	70°C; 0.3xSSC
F	RNA:RNA	< 50	T <sub>F</sub> ; 1xSSC	T <sub>F</sub> ; 1xSSC
G	DNA:DNA	> 50	65°C; 4xSSC -or- 42°C; 4xSSC, 50% formamide	65°C; 1xSSC
H	DNA:DNA	< 50	T <sub>H</sub> ; 4xSSC	T <sub>H</sub> ; 4xSSC
I	DNA:RNA	> 50	67°C; 4xSSC -or- 45°C; 4xSSC, 50% formamide	67°C; 1xSSC
J	DNA:RNA	< 50	T <sub>J</sub> ; 4xSSC	T <sub>J</sub> ; 4xSSC
K	RNA:RNA	> 50	70°C; 4xSSC -or- 50EC; 4xSSC, 50% formamide	67°C; 1xSSC
L	RNA:RNA	< 50	T <sub>L</sub> ; 2xSSC	T <sub>L</sub> ; 2xSSC
M	DNA:DNA	> 50	50°C; 4xSSC -or- 40°C; 6xSSC, 50% formamide	50°C; 2xSSC
N	DNA:DNA	< 50	T <sub>N</sub> ; 6xSSC	T <sub>N</sub> ; 6xSSC
O	DNA:RNA	> 50	55°C; 4xSSC -or- 42°C; 6xSSC, 50% formamide	55°C; 2xSSC
P	DNA:RNA	< 50	T <sub>P</sub> ; 6xSSC	T <sub>P</sub> ; 6xSSC
Q	RNA:RNA	> 50	60°C; 4xSSC -or- 45°C; 6xSSC, 50% formamide	60°C; 2xSSC
R	RNA:RNA	< 50	T <sub>R</sub> ; 4xSSC	T <sub>R</sub> ; 4xSSC

(bp)<sup>1</sup>: The hybrid length is that anticipated for the hybridized region(s) of the hybridizing polynucleotides. When hybridizing a polynucleotide to a target ,



polynucleotide of unknown sequence, the hybrid length is assumed to be that of the hybridizing polynucleotide. When polynucleotides of known sequence are hybridized, the hybrid length can be determined by aligning the sequences of the polynucleotides and identifying the region or regions of optimal sequence complementarity.

Buffer<sup>H</sup>: SSPE (1xSSPE is 0.15M NaCl, 10mM NaH<sub>2</sub>PO<sub>4</sub>, and 1.25mM EDTA, pH 7.4), can be substituted for SSC (1xSSC is 0.15M NaCl and 15mM sodium citrate) in the hybridization and wash buffers; washes are performed for 15 minutes after hybridization is complete.

T<sub>B</sub> through T<sub>R</sub>: The hybridization temperature for hybrids anticipated to be less than 50 base pairs in length should be 5-10°C less than the melting temperature (T<sub>m</sub>) of the hybrid, where T<sub>m</sub> is determined according to the following equations. For hybrids less than 18 base pairs in length, T<sub>m</sub>(EC) = 2(# of A + T bases) + 4(# of G + C bases). For hybrids between 18 and 49 base pairs in length, T<sub>m</sub>(EC) = 81.5 + 16.6(log<sub>10</sub>[Na<sup>+</sup>]) + 0.41(%G+C) - (600/N), where N is the number of bases in the hybrid, and [Na<sup>+</sup>] is the concentration of sodium ions in the hybridization buffer ([Na<sup>+</sup>] for 1xSSC = 0.165 M).

Additional examples of stringency conditions for polynucleotide hybridization are provided in Sambrook *et al.*, 1989, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, chapters 9 and 11, and Ausubel *et al.*, 1995, Current Protocols in Molecular Biology, Eds., John Wiley & Sons, Inc., sections 2.10 and 6.3-6.4, incorporated herein by reference.

In addition to the nucleic acid molecules encoding *Alloiococcus otitidis* polypeptides described above, another aspect of the invention pertains to isolated nucleic acid molecules that are antisense thereto. An "antisense" nucleic acid comprises a nucleotide sequence that is complementary to a "sense" nucleic acid encoding a protein, *e.g.*, complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. Accordingly, an antisense nucleic acid can hydrogen bond to a sense nucleic acid. The antisense nucleic acid can be complementary to an entire *Alloiococcus otitidis* coding strand, or to only a fragment thereof. In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence encoding an *Alloiococcus otitidis* polypeptide.

The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid residues, *e.g.*, the entire coding region of each of the odd numbered sequences set forth in SEQ ID NO: 1 through SEQ ID NO: 105. In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence encoding an *Alloiococcus otitidis* polypeptide. The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids (*i.e.*, also referred to as 5' and 3' untranslated regions).

Given the coding strand sequence encoding the *Alloiococcus otitidis* polypeptides disclosed herein antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of *Alloiococcus otitidis* mRNA, but more preferably is an oligonucleotide which is antisense to only a fragment of the coding or noncoding region of *Alloiococcus otitidis* mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of *Alloiococcus otitidis* mRNA.

An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis and enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (*e.g.*, an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, *e.g.*, phosphorothioate derivatives and acridine substituted nucleotides can be used. Examples of modified nucleotides which can be used to generate the antisense nucleic acid include 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxymethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-

methythio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine.

Alternatively, the antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (*i.e.*, RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

The antisense nucleic acid molecules of the invention are typically administered to a subject or generated *in situ* such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding an *Alloiococcus otitidis* polypeptide to thereby inhibit expression of the polypeptide, *e.g.*, by inhibiting transcription and/or translation. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule which binds to DNA duplexes, through specific interactions in the major groove of the double helix. An example of a route of administration of an antisense nucleic acid molecule of the invention includes direct injection at a tissue site.

Alternatively, an antisense nucleic acid molecule can be modified to target selected cells and then administered systemically. For example, for systemic administration, an antisense molecule can be modified such that it specifically binds to a receptor or an antigen expressed on a selected cell surface, *e.g.*, by linking the antisense nucleic acid molecule to a peptide or an antibody which binds to a cell surface receptor or antigen. The antisense nucleic acid molecule can also be delivered to cells using the vectors described herein.

In yet another embodiment, the antisense nucleic acid molecule of the invention is an  $\alpha$ -anomeric nucleic acid molecule. An  $\alpha$ -anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual  $\gamma$ -units, the strands run parallel to each other (Gaultier *et al.*, 1987). The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (Inoue *et al.*, 1987) or a chimeric RNA-DNA analogue (Inoue *et al.*, 1987).

In still another embodiment, an antisense nucleic acid of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity that are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (*e.g.*, hammerhead ribozymes  
5 described in Haselhoff and Gerlach, 1988) can be used to catalytically cleave *Alloiococcus otitidis* mRNA transcripts to thereby inhibit translation of *Alloiococcus otitidis* mRNA. A ribozyme having specificity for an *Alloiococcus otitidis*-encoding nucleic acid can be designed based upon the nucleotide sequence of an *Alloiococcus otitidis* cDNA disclosed herein. For example, a derivative of a  
10 Tetrahymena L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in an *Alloiococcus otitidis*-encoding mRNA. See, *e.g.*, Cech *et al.* U.S. 4,987,071 and Cech *et al.* U.S. 5,116,742 both incorporated herein in their entirety by reference. Alternatively, *Alloiococcus otitidis* mRNA can be used to select a catalytic RNA  
15 having a specific ribonuclease activity from a pool of RNA molecules. See, *e.g.*, Bartel and Szostak, 1993.

Alternatively *Alloiococcus otitidis* gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region of the *Alloiococcus otitidis* gene (*e.g.*, the *Alloiococcus otitidis* gene promoter and/or  
20 enhancers) to form triple helical structures that prevent transcription of the *Alloiococcus otitidis* gene in target cells. See generally, Helene, 1991; Helene *et al.*, 1992; and Maher, 1992.

*Alloiococcus otitidis* gene expression can also be inhibited using RNA interference (RNAi). This is a technique for post-transcriptional gene silencing  
25 (PTGS), in which target gene activity is specifically abolished with cognate double-stranded RNA (dsRNA). RNAi resembles in many aspects PTGS in plants and has been detected in many invertebrates including trypanosome, hydra, planaria, nematode and fruit fly (*Drosophila melanogaster*). It may be involved in the modulation of transposable element mobilization and antiviral state formation. RNAi  
30 in mammalian systems is disclosed in WO 00/63364, which is incorporated by reference herein in its entirety. Basically, dsRNA of at least about 600 nucleotides, homologous to the target is introduced into the cell and a sequence specific reduction in gene activity is observed.

**C. ALLOIOCOCCUS OTITIDIS POLYPEPTIDES**

In particular embodiments, the present invention provides isolated and  
5 purified *Alloiococcus otitidis* polypeptides. Preferably, an *Alloiococcus otitidis*  
polypeptide of the invention is a recombinant polypeptide. In certain embodiments,  
an *Alloiococcus otitidis* polypeptide of the present invention comprises the amino acid  
sequence that has at least 25% identity to the amino acid sequence of one of the  
even numbered sequences set forth in SEQ ID NO: 2 through SEQ ID NO: 106, a  
10 biological equivalent thereof, or a fragment thereof.

An *Alloiococcus otitidis* polypeptide according to the present invention  
encompasses a polypeptide that comprises: 1) the amino acid sequence shown in  
one of the even numbered sequences set forth in SEQ ID NO: 2 through SEQ ID NO:  
106) functional and non-functional naturally occurring variants or biological  
15 equivalents of *Alloiococcus otitidis* polypeptides of the even numbered sequences set  
forth in SEQ ID NO: 2 through SEQ ID NO: 106 and recombinantly produced variants  
or biological equivalents of *Alloiococcus otitidis* polypeptides set out in SEQ ID NO: 2  
through SEQ ID NO: 106) polypeptides isolated from organisms other than  
*Alloiococcus otitidis* (orthologs of *Alloiococcus otitidis* polypeptides.)

20 A biological equivalent or variant of an *Alloiococcus otitidis* polypeptide  
according to the present invention encompasses 1) a polypeptide isolated from  
*Alloiococcus otitidis*; and 2) a polypeptide that contains substantial homology to an  
*Alloiococcus otitidis* polypeptide.

Biological equivalents or variants of *Alloiococcus otitidis* include both  
25 functional and non-functional *Alloiococcus otitidis* polypeptides. Functional biological  
equivalents or variants are naturally occurring amino acid sequence variants of an  
*Alloiococcus otitidis* polypeptide that maintain the ability to elicit an immunological or  
antigenic response in a subject. Functional variants will typically contain only  
conservative substitutions of one or more amino acids in any one of even numbered  
30 sequences set forth in SEQ ID NO: 2 through SEQ ID NO: 106 or substitution,  
deletion or insertion of non-critical residues in non-critical regions of the polypeptide.

The present invention further provides non-*Alloiococcus otitidis* orthologues of  
*Alloiococcus otitidis* polypeptides. Orthologues of *Alloiococcus otitidis* polypeptides

are polypeptides that are isolated from non-*Alloicoccus otitidis* organisms and possess antigenic capabilities of the *Alloicoccus otitidis* polypeptide. Orthologues of an *Alloicoccus otitidis* polypeptide can readily be identified as comprising an amino acid sequence that is substantially homologous to one of the even numbered sequences set forth in SEQ ID NO: 2 through SEQ ID NO: 106.

Modifications and changes can be made in the structure of a polypeptide of the present invention and still obtain a molecule having *Alloicoccus otitidis* antigenicity. For example, certain amino acids can be substituted for other amino acids in a sequence without appreciable loss of antigenicity. Because it is the interactive capacity and nature of a polypeptide that defines that polypeptide's biological functional activity, certain amino acid sequence substitutions can be made in a polypeptide sequence (or, of course, its underlying DNA coding sequence) and nevertheless obtain a polypeptide with like properties.

In making such changes, the hydropathic index of amino acids can be considered. The importance of the hydropathic amino acid index in conferring interactive biologic function on a polypeptide is generally understood in the art (Kyte & Doolittle, 1982). It is known that certain amino acids can be substituted for other amino acids having a similar hydropathic index or score and still result in a polypeptide with similar biological activity. Each amino acid has been assigned a hydropathic index on the basis of its hydrophobicity and charge characteristics. Those indices are: isoleucine (+4.5); valine (+4.2); leucine (+3.8); phenylalanine (+2.8); cysteine/cystine (+2.5); methionine (+1.9); alanine (+1.8); glycine (-0.4); threonine (-0.7); serine (-0.8); tryptophan (-0.9); tyrosine (-1.3); proline (-1.6); histidine (-3.2); glutamate (-3.5); glutamine (-3.5); aspartate (-3.5); asparagine (-3.5); lysine (-3.9); and arginine (-4.5).

It is believed that the relative hydropathic character of the amino acid residue determines the secondary and tertiary structure of the resultant polypeptide, which in turn defines the interaction of the polypeptide with other molecules, such as enzymes, substrates, receptors, antibodies, antigens, and the like. It is known in the art that an amino acid can be substituted by another amino acid having a similar hydropathic index and still obtain a functionally equivalent polypeptide. In such changes, the substitution of amino acids whose hydropathic indices are within  $\pm 2$  is

preferred, those within  $\pm 1$  are particularly preferred, and those within  $\pm 0.5$  are even more particularly preferred.

Substitution of like amino acids can also be made on the basis of hydrophilicity, particularly where the biologically functional equivalent polypeptide or peptide thereby created is intended for use in immunological embodiments. U.S. Pat. No. 4,554,101, incorporated herein by reference, states that the greatest local average hydrophilicity of a polypeptide, as governed by the hydrophilicity of its adjacent amino acids, correlates with its immunogenicity and antigenicity, *i.e.* with a biological property of the polypeptide.

As detailed in U.S. Pat. No. 4,554,101, the following hydrophilicity values have been assigned to amino acid residues: arginine (+3.0); lysine (+3.0); aspartate (+3.0  $\pm 1$ ); glutamate (+3.0  $\pm 1$ ); serine (+0.3); asparagine (+0.2); glutamine (+0.2); glycine (0); proline (-0.5  $\pm 1$ ); threonine (-0.4); alanine (-0.5); histidine (-0.5); cysteine (-1.0); methionine (-1.3); valine (-1.5); leucine (-1.8); isoleucine (-1.8); tyrosine (-2.3); phenylalanine (-2.5); tryptophan (-3.4). It is understood that an amino acid can be substituted for another having a similar hydrophilicity value and still obtain a biologically equivalent, and in particular, an immunologically equivalent polypeptide. In such changes, the substitution of amino acids whose hydrophilicity values are within  $\pm 2$  is preferred, those which are within  $\pm 1$  are particularly preferred, and those within  $\pm 0.5$  are even more particularly preferred.

As outlined above, amino acid substitutions are generally therefore based on the relative similarity of the amino acid side-chain substituents, for example, their hydrophobicity, hydrophilicity, charge, size, and the like. Exemplary substitutions which take various of the foregoing characteristics into consideration are well known to those of skill in the art and include: arginine and lysine; glutamate and aspartate; serine and threonine; glutamine and asparagine; and valine, leucine and isoleucine (See Table 3, below). The present invention thus contemplates functional or biological equivalents of an *Alloiococcus otitidis* polypeptide as set forth above.

**TABLE 3:**  
**AMINO ACID SUBSTITUTIONS**

<b>Original Residue</b>	<b>Exemplary Residue Substitution</b>
Ala	Gly; Ser
Arg	Lys
Asn	Gln; His
Asp	Glu
Cys	Ser
Gln	Asn
Glu	Asp
Gly	Ala
His	Asn; Gln
Ile	Leu; Val
Leu	Ile; Val
Lys	Arg
Met	Met; Leu; Tyr
Ser	Thr
Thr	Ser
Trp	Tyr
Tyr	Trp; Phe
Val	Ile; Leu

Biological or functional equivalents of a polypeptide are also prepared using site-specific mutagenesis. Site-specific mutagenesis is a technique useful in the preparation of second generation polypeptides, or biologically functional equivalent polypeptides or peptides, derived from the sequences thereof, through specific mutagenesis of the underlying DNA. As noted above, such changes can be desirable where amino acid substitutions are desirable. The technique further provides a capacity to prepare and test sequence variants, for example, incorporating one or more of the foregoing considerations, by introducing one or more nucleotide sequence changes into the DNA. Site-specific mutagenesis allows the production of mutants through the use of specific oligonucleotide sequences which encode the DNA sequence of the desired mutation, as well as a sufficient number of adjacent nucleotides, to provide a primer sequence of sufficient size and sequence complexity to form a stable duplex on both sides of the deletion junction being traversed. Typically, a primer of about 17 to 25 nucleotides in length is preferred, with about 5 to 10 residues on both sides of the site of the alteration of the sequence.

In general, the technique of site-specific mutagenesis is well known in the art. As will be appreciated, the technique typically employs a phage vector, that can exist



in both a single stranded and double stranded form. Typically, site-directed mutagenesis in accordance herewith is performed by first obtaining a single-stranded vector which includes within its sequence a DNA sequence which encodes all or a portion of the *Alloiococcus otitidis* polypeptide sequence selected. An

5 oligonucleotide primer bearing the desired mutated sequence is prepared (*e.g.*, synthetically). This primer is then annealed to the singled-stranded vector, and extended by the use of enzymes such as *Escherichia coli* polymerase I Klenow fragment, in order to complete the synthesis of the mutation-bearing strand. Thus, a heteroduplex is formed wherein one strand encodes the original non-mutated  
10 sequence and the second strand bears the desired mutation. This heteroduplex vector is then used to transform appropriate cells such as *Escherichia coli* cells and clones are selected which include recombinant vectors bearing the mutation. Commercially available kits come with all the reagents necessary, except the oligonucleotide primers.

15 An *Alloiococcus otitidis* polypeptide or polypeptide antigen of the present invention is understood to be any *Alloiococcus otitidis* polypeptide comprising substantial sequence similarity, structural similarity and/or functional similarity to an *Alloiococcus otitidis* polypeptide comprising the amino acid sequence of one of the even numbered sequences set forth in SEQ ID NO: 2 through SEQ ID NO: 106. In  
20 addition, an *Alloiococcus otitidis* polypeptide or polypeptide antigen of the invention is not limited to a particular source. Thus, the invention provides for the general detection and isolation of the polypeptides from a variety of sources.

It is contemplated in the present invention, that an *Alloiococcus otitidis* polypeptide may advantageously be cleaved into fragments for use in further  
25 structural or functional analysis, or in the generation of reagents such as *Alloiococcus otitidis*-related polypeptides and *Alloiococcus otitidis*-specific antibodies. This can be accomplished by treating purified or unpurified *Alloiococcus otitidis* polypeptides with a peptidase such as endoproteinase glu-C (Boehringer, Indianapolis, IN). Treatment with CNBr is another method by which peptide fragments may be produced from  
30 natural *Alloiococcus otitidis* polypeptides. Recombinant techniques also can be used to produce specific fragments of an *Alloiococcus otitidis* polypeptide.

In addition, the inventors also contemplate that compounds sterically similar to a particular *Alloiococcus otitidis* polypeptide antigen, called peptidomimetics, may

be formulated to mimic the key portions of the peptide structure. Peptidemimetics are peptide-containing molecules that mimic elements of protein secondary structure. (See, for example, Johnson *et al.*, 1993.) The underlying rationale behind the use of peptide mimetics is that the peptide backbone of proteins exists chiefly to orient amino acid side chains in such a way as to facilitate molecular interactions, such as those of receptor and ligand.

Successful applications of the peptide mimetic concept have thus far focused on mimetics of  $\beta$ -turns within proteins. Likely  $\beta$ -turn structures, within *Alloiococcus otitidis*, can be predicted by computer-based algorithms as discussed above. Once the component amino acids of the turn are determined, mimetics can be constructed to achieve a similar spatial orientation of the essential elements of the amino acid side chains, as discussed in Johnson *et al.*, 1993.

Fragments of the *Alloiococcus otitidis* polypeptides are also included in the invention. A fragment is a polypeptide having an amino acid sequence that entirely is the same as a part, but not all, of the amino acid sequence. The fragment can comprise, for example, at least 7 or more (e.g., 8, 10, 12, 14, 16, 18, 20 or more) contiguous amino acids of an amino acid sequence selected from one of the even numbered sequences set forth in SEQ ID NO.: 2 through SEQ ID NO.: 106. Fragments may be "freestanding" or comprised within a larger polypeptide of which they form a part or region, most preferably as a single, continuous region. In one embodiment, the fragments include at least one epitope of the mature polypeptide sequence.

"Fusion protein" refers to a protein encoded by two, often unrelated, fused genes or fragments thereof. For example, fusion proteins comprising various portions of constant region of immunoglobulin molecules together with another human protein or part thereof have been described. In many cases, employing an immunoglobulin Fc region as a part of a fusion protein is advantageous for use in therapy and diagnosis resulting in, for example, improved pharmacokinetic properties (see, e.g., EP-A 0232 2621). On the other hand, for some uses it would be desirable to be able to delete the Fc part after the fusion protein has been expressed, detected and purified.

#### D. ALLOIOCOCCUS OTTIDIS POLYNUCLEOTIDE AND POLYPEPTIDE VARIANTS

"Variant" as the term is used herein, is a polynucleotide or polypeptide that differs from a reference polynucleotide or polypeptide respectively, but retains essential properties. A typical variant of a polynucleotide differs in nucleotide sequence from another, reference polynucleotide. Changes in the nucleotide sequence of the variant may or may not alter the amino acid sequence of a polypeptide encoded by the reference polynucleotide. Nucleotide changes may result in amino acid substitutions, additions, deletions, fusions and truncations in the polypeptide encoded by the reference sequence, as discussed below. A typical variant of a polypeptide differs in amino acid sequence from another, reference polypeptide. Generally, differences are limited so that the sequences of the reference polypeptide and the variant are closely similar overall and, in many regions, identical. A variant and reference polypeptide may differ in amino acid sequence by one or more substitutions, additions and deletions in any combination. A substituted or inserted amino acid residue may or may not be one encoded by the genetic code. A variant of a polynucleotide or polypeptide may be a naturally occurring variant such as an allelic variant, or it may be a variant that is not known to occur naturally. Non-naturally occurring variants of polynucleotides and polypeptides may be made by mutagenesis techniques or by direct synthesis.

"Identity," as known in the art, is a relationship between two or more polypeptide sequences or two or more polynucleotide sequences, as determined by comparing the sequences. In the art, "identity" also means the degree of sequence relatedness between polypeptide or polynucleotide sequences, as the case may be, as determined by the match between strings of such sequences. "Identity" can be readily calculated by known methods, including but not limited to those described in (Computational Molecular Biology, Lesk, A. M., ed., Oxford University Press, New York, 1988; Biocomputing: Informatics and Genome Projects, Smith, D. W., ed., Academic Press, New York, 1993; Computer Analysis of Sequence Data, Part I, Griffin, A. M., and Griffin, H. G., eds., Humana Press, New Jersey, 1994; Sequence Analysis in Molecular Biology, von Heinje, G., Academic Press, 1987; and Sequence Analysis Primer, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991; and Carillo, H., and Lipman, D., SIAM J. Applied Math., 48: 1073 (1988). Preferred methods to determine identity are designed to give the largest match

between the sequences tested. Methods to determine identity are codified in publicly available computer programs. Preferred computer program methods to determine identity between two sequences include, but are not limited to, the GCG program package (Devereux, J., *et al* 1984), BLASTP, BLASTN, and FASTA (Altschul, S. F., *et al.*, 1990. The BLASTX program is publicly available from NCBI and other sources (BLAST Manual, Altschul, S., *et al.*, NCBI NLM NIH Bethesda, Md. 20894; Altschul, S., *et al.*, 1990). The well known Smith-Waterman algorithm may also be used to determine identity.

By way of example, a polynucleotide sequence of the present invention may be identical to the reference sequence of one of SEQ ID NO:1 through SEQ ID NO: 105, that is be 100% identical, or it may include up to a certain integer number of nucleotide alterations as compared to the reference sequence. Such alterations are selected from the group consisting of at least one nucleotide deletion, substitution, including transition and transversion, or insertion, and wherein said alterations may occur at the 5' or 3' terminal positions of the reference nucleotide sequence or anywhere between those terminal positions, interspersed either individually among the nucleotides in the reference sequence or in one or more contiguous groups within the reference sequence. The number of nucleotide alterations is determined by multiplying the total number of nucleotides in one of the odd numbered sequences set forth in SEQ ID NO: 1 through SEQ ID NO: 105 by the numerical percent of the respective percent identity (divided by 100) and subtracting that product from said total number of nucleotides in one of the odd numbered sequences set forth in SEQ ID NO: 1 through SEQ ID NO: 105.

For example, the alterations in an isolated *Alloicoccus otitidis* polynucleotide comprise a polynucleotide sequence that has at least 70% identity to the nucleic acid sequence of one of the odd numbered sequences set forth in SEQ ID NO: 1 through SEQ ID NO: 105; a degenerate variant thereof or a fragment thereof, wherein the polynucleotide sequence may include up to  $n_n$  nucleic acid alterations over the entire polynucleotide region of the nucleic acid sequence of any one of the odd numbered sequences set forth in SEQ ID NO: 1 through SEQ ID NO: 105, wherein  $n_n$  is the maximum number of alterations and is calculated by the formula:

$$n_n \leq x_n - (x_n * y),$$

in which  $x_n$  is the total number of nucleic acids of one of SEQ ID NO:1 through SEQ ID NO:105 and  $y$  has a value of 0.70, wherein any non-integer product of  $x_n$  and  $y$  is rounded down to the nearest integer prior to subtracting such product from  $x_n$ . Of course,  $y$  may also have a value of 0.80 for 80%, 0.85 for 85%, 0.90 for 90% 0.95 for 95%, etc.

Similarly, a polypeptide sequence of the present invention may be identical to the reference sequence of any one of even numbered sequences set forth in SEQ ID NO: 2 through SEQ ID NO: 106, that is 100% identical, or it may include up to a certain integer number of amino acid alterations as compared to the reference sequence such that the percentage identity is less than 100%. Such alterations are selected from the group consisting of at least one amino acid deletion, substitution, including conservative and non-conservative substitution, or insertion, and wherein said alterations may occur at the amino- or carboxy-terminal positions of the reference polypeptide sequence or anywhere between those terminal positions, interspersed either individually among the amino acids in the reference sequence or in one or more contiguous groups within the reference sequence. The number of amino acid alterations for a given % identity is determined by multiplying the total number of amino acids in one of the even numbered sequences set forth in SEQ ID NO: 2 through SEQ ID NO: 106 by the numerical percent of the respective percent identity (divided by 100) and then subtracting that product from said total number of amino acids in one of the even numbered sequences set forth in SEQ ID NO: 2 through SEQ ID NO: 106, or:

$$n_a \leq x_a - (x_a \cdot y),$$

wherein  $n_a$  is the number of amino acid alterations,  $x_a$  is the total number of amino acids in one of SEQ ID NO: 2 through SEQ ID NO: 106, and  $y$  is, for instance 0.70 for 70%, 0.80 for 80%, 0.85 for 85% etc., and wherein any non-integer product of  $x_{a.sub.a}$  and  $y$  is rounded down to the nearest integer prior to subtracting it from  $x_a$ .

#### **E. VECTORS, HOST CELLS AND RECOMBINANT *ALLOIOCOCCUS OTITIDIS* POLYPEPTIDES**

In a preferred embodiment, the present invention provides expression vectors comprising ORF polynucleotides that encode *Alloiococcus otitidis* polypeptides. Preferably, the expression vectors of the present invention comprise ORF

polynucleotides that encode *Alloiococcus otitidis* polypeptides comprising the amino acid residue sequence of one of the even numbered sequences set forth in SEQ ID NO: 2 through SEQ ID NO: 106. More preferably, the expression vectors of the present invention comprise a polynucleotide comprising the nucleotide base sequence of the odd numbered sequences set forth in SEQ ID NO: 1 through SEQ ID NO: 105. Even more preferably, the expression vectors of the invention comprise a polynucleotide operatively linked to promoter. Still more preferably, the expression vectors of the invention comprise a polynucleotide operatively linked to a prokaryotic promoter. Alternatively, the expression vectors of the present invention comprise a polynucleotide operatively linked to an enhancer-promoter, that is, an eukaryotic promoter. The expression vectors further comprise a polyadenylation signal that is positioned 3' of the carboxy-terminal amino acid and within a transcriptional unit of the encoded polypeptide.

Expression of proteins in prokaryotes is most often carried out in *Escherichia coli* with vectors containing constitutive or inducible promoters directing the expression of either fusion or non-fusion proteins. Fusion vectors add a number of amino acids to a protein encoded therein, usually to the amino terminus of the recombinant protein. Such fusion vectors typically serve three purposes: 1) to increase expression of recombinant protein; 2) to increase the solubility of the recombinant protein; and 3) to aid in the purification of the recombinant protein by acting as a ligand in affinity purification. Often, in fusion expression vectors, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant protein to enable separation of the recombinant protein from the fusion moiety subsequent to purification of the fusion protein. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase.

Typical fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith and Johnson, 1988), pMAL (New England Biolabs, Beverly; MA) and pRIT5 (Pharmacia, Piscataway, NJ) which fuse glutathione S- transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein.

In one embodiment, the coding sequence of the *Alloiococcus otitidis* polynucleotide is cloned into a pGEX expression vector to create a vector encoding a fusion protein comprising, from the N-terminus to the C-terminus, GST-thrombin cleavage site-*Alloiococcus otitidis* polypeptide. The fusion protein can be purified by

affinity chromatography using glutathione-agarose resin. Recombinant *Alloicoccus otitidis* polypeptide unfused to GST can be recovered by cleavage of the fusion protein with thrombin.

5 Examples of suitable inducible non-fusion *Escherichia coli* expression vectors include pTrc (Amann *et al.*, 1988) and pET 11d (Studier *et al.*, 1990). Target gene expression from the pTrc vector relies on host RNA polymerase transcription from a hybrid trp-lac fusion promoter. Target gene expression from the pET 11d vector relies on transcription from a T7 gn1 0-lac fusion promoter mediated by a coexpressed viral RNA polymerase T7 gnl. This viral polymerase is supplied by host strains BL21 (DE3) or HMS 174(DE3) from a resident prophage harboring a T7 gn1 gene under the transcriptional control of the lacUV 5 promoter.

15 One strategy to maximize recombinant protein expression in *Escherichia coli* is to express the protein in a host bacterium with an impaired capacity to proteolytically cleave the recombinant protein. Another strategy is to alter the nucleic acid sequence of the nucleic acid to be inserted into an expression vector so that the individual codons for each amino acid are those preferentially utilized in *Escherichia coli*. Such alteration of nucleic acid sequences of the invention can be carried out by standard DNA mutagenesis or synthesis techniques.

20 In another embodiment, the *Alloicoccus otitidis* polynucleotide expression vector is a yeast expression vector. Examples of vectors for expression in a yeast such as *S. cerevisiae* include pYepSec I (Baldari, *et al.*, 1987), pMFa (Kurjan and Herskowitz, 1982), pJRY88 (Schultz *et al.*, 1987), and pYES2 (Invitrogen Corporation, San Diego, CA).

25 Alternatively, an *Alloicoccus otitidis* polynucleotide is expressed in insect cells using, for example, baculovirus expression vectors. Baculovirus vectors available for expression of proteins in cultured insect cells (*e.g.*, Sf 9 or Sf 21 cells) include the pAc series (Smith *et al.*, 1983) and the pVL series (Lucklow and Summers, 1989).

30 In yet another embodiment, a nucleic acid of the invention is expressed in mammalian cells using a mammalian expression vector. Examples of mammalian expression vectors include pCDM8 (Seed, 1987) and pMT2PC (Kaufman *et al.*, 1987). When used in mammalian cells, the expression vector's control functions are often provided by viral regulatory elements.

As used herein, a promoter is a region of a DNA molecule typically within about 100 nucleotide pairs in front of (upstream of) the point at which transcription begins (*i.e.*, a transcription start site). That region typically contains several types of DNA sequence elements that are located in similar relative positions in different  
5 genes. As used herein, the term "promoter" includes what is referred to in the art as an upstream promoter region, a promoter region or a promoter of a generalized eukaryotic RNA Polymerase II transcription unit.

Another type of discrete transcription regulatory sequence element is an enhancer. An enhancer provides specificity of time, location and expression level for  
10 a particular encoding region (*e.g.*, gene). A major function of an enhancer is to increase the level of transcription of a coding sequence in a cell that contains one or more transcription factors that bind to that enhancer. Unlike a promoter, an enhancer can function when located at variable distances from transcription start sites so long as a promoter is present.

As used herein, the phrase "enhancer-promoter" means a composite unit that contains both enhancer and promoter elements. An enhancer-promoter is  
15 operatively linked to a coding sequence that encodes at least one gene product. As used herein, the phrase "operatively linked" means that an enhancer-promoter is connected to a coding sequence in such a way that the transcription of that coding  
20 sequence is controlled and regulated by that enhancer-promoter. Means for operatively linking an enhancer-promoter to a coding sequence are well known in the art. As is also well known in the art, the precise orientation and location relative to a coding sequence whose transcription is controlled, is dependent *inter alia* upon the specific nature of the enhancer-promoter. Thus, a TATA box minimal promoter is  
25 typically located from about 25 to about 30 base pairs upstream of a transcription initiation site and an upstream promoter element is typically located from about 100 to about 200 base pairs upstream of a transcription initiation site. In contrast, an enhancer can be located downstream from the initiation site and can be at a considerable distance from that site.

30 An enhancer-promoter used in a vector construct of the present invention can be any enhancer-promoter that drives expression in a cell to be transfected. By employing an enhancer-promoter with well-known properties, the level and pattern of gene product expression can be optimized.



For example, commonly used promoters are derived from polyoma, Adenovirus 2, cytomegalovirus (CMV) and Simian Virus 40 (SV40). For other suitable expression systems for both prokaryotic and eukaryotic cells see chapters 16 and 17 of Sambrook *et al.*, "Molecular Cloning: A Laboratory Manual" 2nd, ed, Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989, incorporated herein by reference.

In another embodiment, the recombinant mammalian expression vector is capable of directing expression of the nucleic acid preferentially in a particular cell type (*e.g.*, tissue-specific regulatory elements are used to express the nucleic acid). Tissue-specific regulatory elements are known in the art. Non-limiting examples of suitable tissue-specific promoters include the albumin promoter (liver-specific; Pinkert *et al.*, 1987), lymphoid-specific promoters (Calame and Eaton, 1988), in particular promoters of T cell receptors (Winoto and Baltimore, 1989) and immunoglobulins (Banerji *et al.*, 1983), Queen and Baltimore (1983), neuron-specific promoters (*e.g.*, the neurofilament promoter; Byrne and Ruddle, 1989), pancreas-specific promoters (Edlund *et al.*, 1985), and mammary gland-specific promoters (*e.g.*, milk whey promoter; U.S. 4, 873,316 and EP 264,166). Developmentally-regulated promoters are also encompassed, for example the murine hox promoters (Kessel and Gruss, 1990) and the  $\alpha$ -fetoprotein promoter (Campes and Tilghman, 1989).

The invention further provides a recombinant expression vector comprising a DNA molecule encoding an *Alloiococcus otitidis* polypeptide cloned into the expression vector in an antisense orientation. That is, the DNA molecule is operatively linked to a regulatory sequence in a manner which allows for expression (by transcription of the DNA molecule) of an RNA molecule which is antisense to *Alloiococcus otitidis* mRNA. Regulatory sequences operatively linked to a nucleic acid cloned in the antisense orientation can be chosen which direct the continuous expression of the antisense RNA molecule in a variety of cell types, for instance viral promoters and/or enhancers, or regulatory sequences can be chosen which direct constitutive, tissue specific or cell type specific expression of antisense RNA. The antisense expression vector can be in the form of a recombinant plasmid, phagemid or attenuated virus in which antisense nucleic acids are produced under the control of a high efficiency regulatory region, the activity of which can be determined by the cell type into which the vector is introduced.

Another aspect of the invention pertains to host cells into which a recombinant expression vector of the invention has been introduced. The terms "host cell" and "recombinant host cell" are used interchangeably herein. It is understood that such terms refer not only to the particular subject cell but also to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein. A host cell can be any prokaryotic or eukaryotic cell. For example, an *Alloiococcus otitidis* polypeptide can be expressed in bacterial cells such as *Escherichia coli*, insect cells, yeast or mammalian cells (such as Chinese hamster ovary cells (CHO), NIH3T3, PER C6, NSO, VERO or COS cells). Other suitable host cells are known to those skilled in the art.

Vector DNA can be introduced into prokaryotic or eukaryotic cells via conventional transformation, infection or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid (e.g., DNA) into a host cell, including calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, lipofection, protoplast fusion, direct microinfection. Another recognized technique for introducing DNA into a host cell is "infection", such as by adenovirus infection or electroporation. Suitable methods for transforming, infecting or transfecting host cells can be found in Sambrook, *et al.* ("Molecular Cloning: A Laboratory Manual" 2nd ed, Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989), and other laboratory manuals.

The most widely used method is transfection mediated by either calcium phosphate or DEAE-dextran. Although the mechanism remains unclear, it is believed that the transfected DNA enters the cytoplasm of the cell by endocytosis and is transported to the nucleus. Depending on the cell type, up to 90% of a population of cultured cells can be transfected at any one time. Because of its high efficiency, transfection mediated by calcium phosphate or DEAE-dextran is the method of choice for experiments that require transient expression of the foreign DNA in large numbers of cells. Calcium phosphate-mediated transfection is also used to establish cell lines that integrate copies of the foreign DNA, which are usually arranged in head-to-tail tandem arrays into the host cell genome.

In the protoplast fusion method, protoplasts derived from bacteria carrying high numbers of copies of plasmid of interest are mixed directly with cultured mammalian cells. After fusion of the cell membranes (usually with polyethylene glycol), the contents of the bacteria are delivered into the cytoplasm of the mammalian cells and the plasmid DNA is transported to the nucleus. Protoplast fusion is not as efficient as transfection for many of the cell lines that are commonly used for transient expression assays, but it is useful for cell lines in which endocytosis of DNA occurs inefficiently. Protoplast fusion frequently yields multiple copies of the plasmid DNA tandemly integrated into the host chromosome.

The application of brief, high-voltage electric pulses (electroporation) to a variety of mammalian and plant cells leads to the formation of nanometer-sized pores in the plasma membrane. DNA is taken directly into the cell cytoplasm either through these pores or as a consequence of the redistribution of membrane components that accompanies closure of the pores. Electroporation can be extremely efficient and can be used both for transient expression of cloned genes and for establishment of cell lines that carry integrated copies of the gene of interest. Electroporation, in contrast to calcium phosphate-mediated transfection and protoplast fusion, frequently gives rise to cell lines that carry one, or at most a few, integrated copies of the foreign DNA.

Liposome transfection involves encapsulation of DNA and RNA within liposomes, followed by fusion of the liposomes with the cell membrane. The mechanism of how DNA is delivered into the cell is unclear, but transfection efficiencies can be as high as 90%.

Direct microinjection of a DNA molecule into nuclei has the advantage of not exposing DNA to cellular compartments such as low-pH endosomes. Microinjection therefore used primarily as a method to establish lines of cells that carry integrated copies of the DNA of interest.

The use of adenovirus as a vector for cell transfection is well known in the art. Adenovirus vector-mediated cell transfection has been reported for various cells (Stratford-Perricaudet, *et al.* 1992).

A host cell of the invention, such as a prokaryotic or eukaryotic host cell in culture, is used to produce (*i.e.*, express) an *Alloiococcus otitidis* polypeptide. Accordingly, the invention further provides methods for producing an *Alloiococcus*

*otitidis* polypeptide using the host cells of the invention. In one embodiment, the method comprises culturing the host cell of invention (into which a recombinant expression vector encoding an *Alloiococcus otitidis* polypeptide has been introduced) in a suitable medium until the *Alloiococcus otitidis* polypeptide is produced. In  
5 another embodiment, the method further comprises isolating the *Alloiococcus otitidis* polypeptide from the medium or the host cell.

A coding sequence of an expression vector is operatively linked to a transcription-terminating region. RNA polymerase transcribes an encoding DNA sequence through a site where polyadenylation occurs. Typically, DNA sequences  
10 located a few hundred base pairs downstream of the polyadenylation site serve to terminate transcription. Those DNA sequences are referred to herein as transcription-termination regions. Those regions are required for efficient polyadenylation of transcribed messenger RNA (mRNA). Transcription-terminating regions are well known in the art. A preferred transcription-terminating region used in  
15 an adenovirus vector construct of the present invention comprises a polyadenylation signal of SV40 or the protamine gene.

An expression vector comprises a polynucleotide that encodes an *Alloiococcus otitidis* polypeptide. Such a polypeptide is meant to include a sequence of nucleotide bases encoding an *Alloiococcus otitidis* polypeptide sufficient in length  
20 to distinguish the segment from a polynucleotide segment encoding a non-*Alloiococcus otitidis* polypeptide. A polypeptide of the invention can also encode biologically functional polypeptides or peptides which have variant amino acid sequences, such as with changes selected based on considerations such as the relative hydropathic score of the amino acids being exchanged. These variant  
25 sequences are those isolated from natural sources or induced in the sequences disclosed herein using a mutagenic procedure such as site-directed mutagenesis.

Preferably, an expression vector of the present invention comprises a polynucleotide that encodes a polypeptide comprising the amino acid residue sequence of one of the even numbered sequences set forth in SEQ ID NO: 2 through  
30 SEQ ID NO: 4036. An expression vector can include an *Alloiococcus otitidis* polypeptide coding region itself of any of the *Alloiococcus otitidis* polypeptides noted above or it can contain coding regions bearing selected alterations or modifications in the basic coding region of such an *Alloiococcus otitidis* polypeptide. Alternatively,

such vectors or fragments can also encode larger polypeptides or polypeptides which nevertheless include the basic coding region. In any event, it should be appreciated that due to codon redundancy as well as biological functional equivalence, this aspect of the invention is not limited to the particular DNA molecules corresponding to the polypeptide sequences noted above.

Exemplary vectors include the mammalian expression vectors of the pCMV family including pCMV6b and pCMV6c (Chiron Corp., Emeryville CA.). In certain cases, and specifically in the case of these individual mammalian expression vectors, the resulting constructs can require co-transfection with a vector containing a selectable marker such as pSV2neo. Via co-transfection into a dihydrofolate reductase-deficient Chinese hamster ovary cell line, such as DG44, clones expressing *Alloiococcus otitidis* polypeptides by virtue of DNA incorporated into such expression vectors can be detected.

A DNA molecule of the present invention can be incorporated into a vector by a number of techniques that are well known in the art. For instance, the vector pUC18 has been demonstrated to be of particular value in cloning and expression of genes. Likewise, the related vectors M13mp18 and M13mp19 can also be used in certain embodiments of the invention, in particular, in performing dideoxy sequencing.

An expression vector of the present invention is useful both as a means for preparing quantities of the *Alloiococcus otitidis* polypeptide-encoding DNA itself, and as a means for preparing the encoded polypeptide and peptides. It is contemplated that where *Alloiococcus otitidis* polypeptides of the invention are made by recombinant means, one can employ either prokaryotic or eukaryotic expression vectors as shuttle systems. In another aspect, the recombinant host cells of the present invention are prokaryotic host cells. Preferably, the recombinant host cells of the invention are bacterial cells of the DH5 $\alpha$  strain of *Escherichia coli*. In general, prokaryotes are preferred for the initial cloning of DNA sequences and constructing the vectors useful in the invention. For example, *Escherichia coli* K12 strains can be particularly useful. Other microbial strains that can be used include *Escherichia coli* B, *Escherichia coli* W3110 (ATCC No. 273325) and *Escherichia. coli*<sub>x</sub>1976 (ATCC No. 31537). *Bacilli* such as *Bacillus subtilis*, or other enterobacteriaceae such as *Salmonella typhimurium* or other *Salmonella* species or *Serratia marcesans*, and

various pseudomonas species can be used. These examples are, of course, intended to be illustrative rather than limiting.

In general, plasmid vectors containing replicon and control sequences that are derived from species compatible with the host cell are used in connection with these hosts. The vector ordinarily carries a replication site, as well as marking sequences that are capable of providing phenotypic selection in transformed cells. For example, *Escherichia coli* can be transformed using pBR322, a plasmid derived from an *Escherichia coli* species (Bolivar, *et al.* 1977). pBR322 contains genes for ampicillin and tetracycline resistance and thus provides easy means for identifying transformed cells. The pBR322 plasmid, or other microbial plasmid or phage, must also contain, or be modified to contain, promoters which can be used by the microbial organism for expression of its own polypeptides.

Those promoters most commonly used in recombinant DNA construction include the  $\beta$ -lactamase (penicillinase) and lactose promoter systems (Chang, *et al.* 1978; Itakura, *et al.* 1977, Goeddel, *et al.* 1979; Goeddel, *et al.* 1980) and a tryptophan (TRP) promoter system (EP 0036776; Siebwenlist *et al.* 1980). While these are the most commonly used, other microbial promoters have been discovered and utilized, and details concerning their nucleotide sequences have been published, enabling a skilled worker to introduce functional promoters into plasmid vectors (Siebwenlist, *et al.* 1980).

In addition to prokaryotes, eukaryotic microbes such as yeast can also be used. *Saccharomyces cerevisiae* or common baker's yeast is the most commonly used among eukaryotic microorganisms, although a number of other strains are commonly available. For expression in *Saccharomyces*, the plasmid YRp7, for example, is commonly used (Stinchcomb, *et al.* 1979; Kingsman, *et al.* 1979; Tschemper, *et al.* 1980). This plasmid already contains the *trp1* gene that provides a selection marker for a mutant strain of yeast lacking the ability to grow in tryptophan, for example ATCC No. 44076 or PEP4-1 (Jones, 1977). The presence of the *trp1* lesion as a characteristic of the yeast host cell genome then provides an effective environment for detecting transformation by growth in the absence of tryptophan.

Suitable promoter sequences in yeast vectors include the promoters for 3-phosphoglycerate kinase (PGK) (Hitzeman, *et al.* 1980) or other glycolytic enzymes (Hess, *et al.* 1968; Holland, *et al.* 1978) such as enolase, glyceraldehyde-3-

phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phosphoglucose isomerase, and glucokinase. In constructing suitable expression plasmids, the termination sequences  
5 associated with these genes are also introduced into the expression vector downstream from the sequences to be expressed to provide polyadenylation of the mRNA and termination. Other promoters, which have the additional advantage of transcription controlled by growth conditions are the promoter region for alcohol dehydrogenase 2, isocytochrome C, acid phosphatase, degradative enzymes  
10 associated with nitrogen metabolism, and the aforementioned glyceraldehyde-3-phosphate dehydrogenase, and enzymes responsible for maltose and galactose utilization. Any plasmid vector containing a yeast-compatible promoter, origin of replication, and termination sequences is suitable.

In addition to microorganisms, cultures of cells derived from multicellular  
15 organisms can also be used as hosts. In principle, any such cell culture is workable, whether from vertebrate or invertebrate culture. However, interest has been greatest in vertebrate cells, and propagation of vertebrate cells in culture (tissue culture) has become a routine procedure in recent years. Examples of such useful host cell lines are AtT-20, VERO, HeLa, NSO, PER C6, Chinese hamster ovary (CHO) cell lines,  
20 W138, BHK, COSM6, COS-7, 293, VERO and MDCK cell lines. Expression vectors for such cells ordinarily include (if necessary) an origin of replication, a promoter located upstream of the gene to be expressed, along with any necessary ribosome binding sites, RNA splice sites, polyadenylation site, and transcriptional terminator sequences.

25 Where expression of recombinant *Alloicoccus otitidis* polypeptides is desired and a eukaryotic host is contemplated, it is most desirable to employ a vector, such as a plasmid, that incorporates a eukaryotic origin of replication. Additionally, for the purposes of expression in eukaryotic systems, one desires to position the *Alloicoccus otitidis* encoding sequence adjacent to and under the control of an  
30 effective eukaryotic promoter such as promoters used in combination with Chinese hamster ovary cells (CHO). To bring a coding sequence under control of a promoter, whether it is eukaryotic or prokaryotic, what is generally needed is to position the 5' end of the translation initiation side of the proper translational reading frame of the

polypeptide between about 1 and about 50 nucleotides 3' of or downstream with respect to the promoter chosen. Furthermore, where eukaryotic expression is anticipated, one would typically desire to incorporate an appropriate polyadenylation site into the transcriptional unit that includes the *Alloiococcus otitidis* polypeptide.

5 A transfected cell can be prokaryotic or eukaryotic. Preferably, the host cells of the invention are prokaryotic host cells. Where it is of interest to produce an *Alloiococcus otitidis* polypeptide, cultured prokaryotic host cells are of particular interest.

10 In yet another embodiment, the present invention contemplates a process or method of preparing *Alloiococcus otitidis* polypeptides comprising transfecting, transforming or infecting cells with a polynucleotide that encodes an *Alloiococcus otitidis* polypeptide to produce transformed host cells; and maintaining the transformed host cells under biological conditions sufficient for expression of the polypeptide. Preferably, the transformed host cells are prokaryotic cells.

15 Alternatively, the host cells are eukaryotic cells. More preferably, the prokaryotic cells are bacterial cells of the DH5 $\alpha$  strain of *Escherichia coli*. Even more preferably, the polynucleotide transfected into the transformed cells comprises the nucleic acid sequence of one of the odd numbered sequences set forth in SEQ ID NO: 1 through SEQ ID NO: 105. Additionally, transfection is accomplished using an expression  
20 vector disclosed above. A host cell used in the process is capable of expressing a functional, recombinant *Alloiococcus otitidis* polypeptide.

Following transfection, the cell is maintained under culture conditions for a period of time sufficient for expression of an *Alloiococcus otitidis* polypeptide. Culture conditions are well known in the art and include ionic composition and concentration,  
25 temperature, pH and the like. Typically, transfected cells are maintained under culture conditions in a culture medium. Suitable media for various cell types are well known in the art. In a preferred embodiment, temperature is from about 20°C to about 50°C, more preferably from about 30°C to about 40°C and, even more preferably about 37°C.

30 The pH is preferably from about a value of 6.0 to a value of about 8.0, more preferably from about a value of about 6.8 to a value of about 7.8 and, most preferably about 7.4. Osmolality is preferably from about 200 milliosmols per liter (mosm/L) to about 400 mosm/l and, more preferably from about 290 mosm/L to



about 310 mosm/L. Other biological conditions needed for transfection and expression of an encoded protein are well known in the art.

Transfected cells are maintained for a period of time sufficient for expression of an *Alloiococcus otitidis* polypeptide. A suitable time depends *inter alia* upon the cell type used and is readily determinable by a skilled artisan. Typically, maintenance time is from about 2 to about 14 days.

Recombinant *Alloiococcus otitidis* polypeptide is recovered or collected either from the transfected cells or the medium in which those cells are cultured. Recovery comprises isolating and purifying the *Alloiococcus otitidis* polypeptide. Isolation and purification techniques for polypeptides are well known in the art and include such procedures as precipitation, filtration, chromatography, electrophoresis and the like.

#### F. ANTIBODIES IMMUNOREACTIVE WITH *ALLOIOCOCCUS OTITIDIS* POLYPEPTIDES

In still another embodiment, the present invention provides antibodies immunoreactive with *Alloiococcus otitidis* polypeptides. Preferably, the antibodies of the invention are monoclonal antibodies. Additionally, the *Alloiococcus otitidis* polypeptides comprise the amino acid residue sequence of one of the even numbered sequences set forth in SEQ ID NO: 2 through SEQ ID NO: 106. Means for preparing and characterizing antibodies are well known in the art (See, e.g., Antibodies "A Laboratory Manual", E. Howell and D. Lane, Cold Spring Harbor Laboratory, 1988). Polyclonal antisera is obtained by bleeding an immunized animal into a glass or plastic container, incubating the blood at 25°C for one hour, followed by incubating at 4°C for 2-18 hours. The serum is then recovered by centrifugation.

Briefly, a polyclonal antibody is prepared by immunizing an animal with an immunogen comprising a polypeptide or polynucleotide of the present invention, and collecting antisera from that immunized animal. A wide range of animal species can be used for the production of antisera. Typically an animal used for production of anti-antisera is a rabbit, a mouse, a rat, a hamster or a guinea pig. Because of the relatively large blood volume of rabbits, a rabbit is a preferred choice for production of polyclonal antibodies.

As is well known in the art, a given polypeptide or polynucleotide may vary in its immunogenicity. It is often necessary therefore to couple the immunogen (e.g., a polypeptide or polynucleotide) of the present invention with a carrier. Exemplary and

preferred carriers are keyhole limpet hemocyanin (KLH) and bovine serum albumin (BSA). Other albumins such as ovalbumin, mouse serum albumin or rabbit serum albumin can also be used as carriers.

Means for conjugating a polypeptide or a polynucleotide to a carrier protein are well known in the art and include glutaraldehyde, m-maleimidobencoyl-N-hydroxysuccinimide ester, carbodiimide and bis-biazotized benzidine.

As is also well known in the art, immunogenicity to a particular immunogen can be enhanced by the use of non-specific stimulators of the immune response known as adjuvants. Exemplary and preferred adjuvants include complete Freund's adjuvant, incomplete Freund's adjuvants, cholera toxin (e.g. mutant cholera toxin E29H; *see* published International Patent Application WO 00/18434), and aluminum hydroxide adjuvant.

The amount of immunogen used for the production of polyclonal antibodies depends upon the nature of the immunogen as well as the animal used for immunization. A variety of routes can be used to administer the immunogen (subcutaneous, intramuscular, intradermal, intravenous and intraperitoneal). The production of polyclonal antibodies is monitored by sampling blood from the immunized animal at various points following immunization. When a desired level of immunogenicity is obtained, the immunized animal can be bled and the serum isolated and stored.

In another aspect, the present invention contemplates a process of producing an antibody immunoreactive with an *Alloiococcus otitidis* polypeptide comprising the steps of (a) transfecting recombinant host cells with a polynucleotide that encodes an *Alloiococcus otitidis* polypeptide; (b) culturing the host cells under conditions sufficient for expression of the polypeptide; (c) recovering the polypeptides; and (d) preparing the antibodies to the polypeptides. Preferably, the host cell is transfected with the polynucleotide of one of the odd numbered sequences set forth in SEQ ID NO: 1 through SEQ ID NO: 4035. Even more preferably, the present invention provides antibodies prepared according to the process described above.

A monoclonal antibody of the present invention can be readily prepared through use of well-known techniques such as those exemplified in U.S. Pat. No. 4,196,265, herein incorporated by reference. Typically, a technique involves first immunizing a suitable animal with a selected antigen (e.g., a polypeptide or

polynucleotide of the present invention) in a manner sufficient to provide an immune response. Rodents such as mice and rats are preferred animals. Spleen cells from the immunized animal are then fused with cells of an immortal myeloma cell. Where the immunized animal is a mouse, a preferred myeloma cell is a murine NS-1 myeloma cell.

The fused spleen/myeloma cells are cultured in a selective medium to select fused spleen/myeloma cells from the parental cells. Fused cells are separated from the mixture of non-fused parental cells, *e.g.*, by the addition of agents that block the *de novo* synthesis of nucleotides in the tissue culture media. Exemplary and preferred agents are aminopterin, methotrexate, and azaserine. Aminopterin and methotrexate block *de novo* synthesis of both purines and pyrimidines, whereas azaserine blocks only purine synthesis. Where aminopterin or methotrexate is used, the media is supplemented with hypoxanthine and thymidine as a source of nucleotides. Where azaserine is used, the media is supplemented with hypoxanthine.

This culturing provides a population of hybridomas from which specific hybridomas are selected. Typically, selection of hybridomas is performed by culturing the cells by single-clone dilution in microtiter plates, followed by testing the individual clonal supernatants for reactivity with an antigen-polypeptide. The selected clones can then be propagated indefinitely to provide the monoclonal antibody.

By way of specific example, to produce an antibody of the present invention, mice are injected intraperitoneally with between about 1-200  $\mu$ g of an antigen comprising a polypeptide of the present invention. B lymphocyte cells are stimulated to grow by injecting the antigen in association with an adjuvant such as complete Freund's adjuvant (CFA; a non-specific stimulator of the immune response containing killed *Mycobacterium tuberculosis*). At some time (*e.g.*, at least two weeks) after the first injection, mice are boosted by injection with a second dose of the antigen mixed with incomplete Freund's adjuvant (IFA; lacks the killed mycobacterium of CFA).

A few weeks after the second injection, mice are tail bled and the sera titrated by immunoprecipitation against radiolabeled antigen. Preferably, the process of boosting and titrating is repeated until a suitable titer is achieved. The spleen of the mouse with the highest titer is removed and the spleen lymphocytes are obtained by

homogenizing the spleen with a syringe. Typically, a spleen from an immunized mouse contains approximately  $5 \times 10^7$  to  $2 \times 10^8$  lymphocytes.

5 Mutant lymphocyte cells known as myeloma cells are obtained from laboratory animals in which such cells have been induced to grow by a variety of well-known methods. Myeloma cells lack the salvage pathway of nucleotide biosynthesis. Because myeloma cells are tumor cells, they can be propagated indefinitely in tissue culture, and are thus denominated immortal. Numerous cultured cell lines of myeloma cells from mice and rats, such as murine NS-1 myeloma cells, have been established.

10 Myeloma cells are combined under conditions appropriate to foster fusion with the normal antibody-producing cells from the spleen of the mouse or rat injected with the antigen/polypeptide of the present invention. Fusion conditions include, for example, the presence of polyethylene glycol. The resulting fused cells are hybridoma cells. Like myeloma cells, hybridoma cells grow indefinitely in culture.

15 Hybridoma cells are separated from unfused myeloma cells by culturing in a selection medium such as HAT media (hypoxanthine, aminopterin, thymidine). Unfused myeloma cells lack the enzymes necessary to synthesize nucleotides from the salvage pathway because they are killed in the presence of aminopterin, methotrexate, or azaserine. Unfused lymphocytes also do not continue to grow in tissue culture. Thus, only cells that have successfully fused (hybridoma cells) can grow in the selection media.

Each of the surviving hybridoma cells produces a single antibody. These cells are then screened for the production of the specific antibody immunoreactive with an antigen/polypeptide of the present invention. Single cell hybridomas are isolated by limiting dilutions of the hybridomas. The hybridomas are serially diluted many times and, after the dilutions are allowed to grow, the supernatant is tested for the presence of the monoclonal antibody. The clones producing that antibody are then cultured in large amounts to produce an antibody of the present invention in convenient quantity.

30 By use of a monoclonal antibody of the present invention, specific polypeptides and polynucleotide of the invention are identified as antigens. Once identified, those polypeptides and polynucleotide are isolated and purified by techniques such as antibody-affinity chromatography. In antibody-affinity

chromatography, a monoclonal antibody is bound to a solid substrate and exposed to a solution containing the desired antigen. The antigen is removed from the solution through an immunospecific reaction with the bound antibody. The polypeptide or polynucleotide is then easily removed from the substrate and purified.

5        Additionally, examples of methods and reagents particularly amenable for use in generating and screening antibody display library can be found in, for example, U.S. 5,223,409; WO 92/18619; WO 91/17271; WO 92/20791; WO 92/15679; WO 93/01288; WO 92/01047; WO 92/09690; WO 90/02809, which are incorporated herein in their entirety by reference.

10        Additionally, recombinant anti-*Alloiococcus otitidis* antibodies, such as chimeric and humanized monoclonal antibodies, comprising both human and non-human fragments, which are made using standard recombinant DNA techniques, are within the scope of the invention. Such chimeric and humanized monoclonal antibodies are produced by recombinant DNA techniques known in the art, for  
15        example using methods described in PCT/US86/02269; EP 184,187; EP 171,496; EP 173,494; WO 86/01533; U.S. 4,816,567; and EP 125,023.

      An anti-*Alloiococcus otitidis* antibody (*e.g.*, monoclonal antibody) is used to isolate *Alloiococcus otitidis* polypeptides by standard techniques, such as affinity chromatography or immunoprecipitation. An anti-*Alloiococcus otitidis* antibody  
20        facilitates the purification of a natural *Alloiococcus otitidis* polypeptide from cells and recombinantly produced *Alloiococcus otitidis* polypeptides expressed in host cells. Moreover, an anti-*Alloiococcus otitidis* antibody is used to detect *Alloiococcus otitidis* polypeptide (*e.g.*, in a cellular lysate or cell supernatant) in order to evaluate the abundance of the *Alloiococcus otitidis* polypeptide. The detection of circulating  
25        fragments of an *Alloiococcus otitidis* polypeptide is used to identify *Alloiococcus otitidis* polypeptide turnover in a subject. Anti-*Alloiococcus otitidis* antibodies are used diagnostically to monitor protein levels in tissue as part of a clinical testing procedure, *e.g.*, to, for example, determine the efficacy of a given treatment regimen. Detection is facilitated by coupling (*i.e.*, physically linking) the antibody to a  
30        detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, P-galactosidase, or

acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and acquorin, and examples of suitable radioactive material include  $^{125}\text{I}$ ,  $^{131}\text{I}$ ,  $^{15}\text{S}$  or  $^3\text{H}$ .

#### G. PHARMACEUTICAL COMPOSITIONS

In certain embodiments, the present invention provides pharmaceutical compositions comprising compounds that inhibit the activities of *Alloicoccus otitidis* polypeptides, and physiologically acceptable carriers. Compounds that inhibit the activities of *Alloicoccus otitidis* polypeptides, which are essential for the proliferation of the bacteria, are identified using one or more assay systems set forth in Examples 5-38. More preferably, the pharmaceutical compositions comprise one or more compounds that inhibit the activities of *Alloicoccus otitidis* polypeptides comprising the amino acid residue sequence of one or more of the even numbered sequences set forth in SEQ ID NO: 2 through SEQ ID NO: 106. In other embodiments, the pharmaceutical compositions of the invention comprise antisense polynucleotides of polynucleotides selected from one of the odd numbered sequences set forth in Seq. ID NO. 1 to Seq. ID No. 105, and physiologically acceptable carriers.

Various tests are to be used to assess the *in vitro* and *in vivo* efficacy of antimicrobial and pharmaceutical compounds that inhibit the activities of *Alloicoccus otitidis* polypeptides, and these are set forth in detail in Examples 5 through 38. For example, an *in vitro* activity of the compounds may be assayed by incubating together a mixture of *Alloicoccus otitidis* or other heterologous bacterial cells such as *E. coli* cells expressing *Alloicoccus otitidis* polypeptides set forth in one of the even numbered sequences from Seq. ID No. 2 to Seq. ID No. 106, and then measuring the activity of the polypeptide using one or more of the assay systems detailed in Example 5 through 38.

The *Alloicoccus otitidis* polynucleotides, polypeptides, compounds that modulate the activity of an *Alloicoccus otitidis* polypeptides, and anti-*Alloicoccus*

otitis antibodies (also referred to herein as "active compounds") of the invention can be incorporated into pharmaceutical compositions suitable for administration to a host or subject, *e.g.*, a human. Such compositions typically comprise the nucleic acid molecule, protein, antimicrobial compound, or antibody and a pharmaceutically acceptable carrier. As used herein the language "pharmaceutically acceptable carrier" is intended to include any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. The use of such media and agents for pharmaceutically active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active compound, such media can be used in the compositions of the invention. Supplementary active compounds can also be incorporated into the compositions.

A pharmaceutical of the invention is formulated to be compatible with its intended route of administration. Examples of routes of administration include parenteral, (*e.g.*, intravenous, intradermal, subcutaneous, intraperitoneal), transmucosal (*e.g.*, oral, rectal, intranasal, vaginal, respiratory), and transdermal (topical). Solutions or suspensions used for parenteral, intradermal, or subcutaneous application can include the following components: a sterile diluent such as water for injection, saline solution, fixed oils, polyethylene glycols, glycerine, propylene glycol or other synthetic solvents; antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfite; chelating agents such as ethylenediaminetetraacetic acid; buffers such as acetates, citrates or phosphates and agents for the adjustment of tonicity such as sodium chloride or dextrose. pH can be adjusted with acids or bases, such as hydrochloric acid or sodium hydroxide. The parenteral preparation can be enclosed in ampoules, disposable syringes or multiple dose vials made of glass or plastic.

Pharmaceutical compositions suitable for injectable use include sterile aqueous solutions (where water-soluble) or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersion. For intravenous administration, suitable carriers include physiological saline, bacteriostatic water, Cremophor EL™ (BASF, Parsippany, NJ) or phosphate buffered saline (PBS). In all cases, the composition must be sterile and should be fluid to the extent that easy syringability exists. It must be stable under the conditions of

can be included as part of the composition. The tablets, pills, capsules, troches and the like can contain any of the following ingredients, or compounds of a similar nature: a binder such as microcrystalline cellulose, gum tragacanth or gelatin; an excipient such as starch or lactose, a disintegrating agent such as alginic acid, Primogel, or corn starch; a lubricant such as magnesium stearate or Sterotes; a glidant such as colloidal silicon dioxide; a sweetening agent such as sucrose or saccharin; or a flavoring agent such as peppermint, methyl salicylate, or orange flavoring.

For administration by inhalation, the compounds are delivered in the form of an aerosol spray from pressured container or dispenser that contains a suitable propellant, *e.g.*, a gas such as carbon dioxide, or a nebulizer. Systemic administration can also be by transmucosal or transdermal means. For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration, detergents, bile salts, and fusidic acid derivatives. Transmucosal administration can be accomplished through the use of nasal sprays or suppositories. For transdermal administration, the active compounds are formulated into ointments, salves, gels, or creams as generally known in the art.

The compounds can also be prepared in the form of suppositories (*e.g.*, with conventional suppository bases such as cocoa butter and other glycerides) or retention enemas for rectal delivery.

In one embodiment, the active compounds are prepared with carriers that will protect the compound against rapid elimination from the body, such as a controlled release formulation, including implants and microencapsulated delivery systems.

Biodegradable, biocompatible polymers can be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid. Methods

#### H. DIAGNOSTIC ASSAYS

The invention also provides methods for detecting the presence of an *Alloiococcus otitidis* polypeptide or *Alloiococcus otitidis* polynucleotide, or fragment thereof, in a biological sample. The method involves contacting the biological sample



with a compound or an agent capable of detecting an *Alloiococcus otitidis* polypeptide or mRNA such that the presence of the *Alloiococcus otitidis* polypeptide/encoding nucleic acid molecule is detected in the biological sample. A preferred agent for detecting *Alloiococcus otitidis* mRNA or DNA is a labeled or labelable oligonucleotide probe capable of hybridizing to *Alloiococcus otitidis* mRNA or DNA. The nucleic acid probe can be, for example, a full-length *Alloiococcus otitidis* polynucleotide of one of the odd numbered sequences set forth in SEQ ID NO: 1 through SEQ ID NO: 105, a complement thereof, or a fragment thereof, such as an oligonucleotide of at least 15, 30, 50, 100, 250 or 500 nucleotides in length and sufficient to specifically hybridize under stringent conditions to *Alloiococcus otitidis* mRNA or DNA. Alternatively, the sample can be contacted with an oligonucleotide primer of an *Alloiococcus otitidis* polynucleotide of SEQ ID NO: 1 through SEQ ID :105, a complement thereof, or a fragment thereof, in the presence of nucleotides and a polymerase, under conditions permitting primer extension.

A preferred agent for detecting *Alloiococcus otitidis* polypeptide is a labeled or labelable antibody capable of binding to an *Alloiococcus otitidis* polypeptide. Antibodies can be polyclonal, or more preferably, monoclonal. An intact antibody, or a fragment thereof (*e.g.*, Fab or F(ab')<sub>2</sub>) can be used. The term "labeled or labelable," with regard to the probe or antibody, is intended to encompass direct labeling of the probe or antibody by coupling (*i.e.*, physically linking) a detectable substance to the probe or antibody, as well as indirect labeling of the probe or antibody by reactivity with another reagent that is directly labeled. Examples of indirect labeling include detection of a primary antibody using a fluorescently labeled secondary antibody and end-labeling of a DNA probe with biotin such that it can be detected with fluorescently labeled streptavidin. The term "biological sample" is intended to include tissues, cells and biological fluids isolated from a subject, as well as tissues, cells and fluids present within a subject. That is, the detection method of the invention can be used to detect *Alloiococcus otitidis* mRNA, DNA or protein in a biological sample *in vitro* as well as *in vivo*. For example, *in vitro* techniques for detection of *Alloiococcus otitidis* mRNA include Northern hybridizations and *in situ* hybridizations. *In vitro* techniques for detection of *Alloiococcus otitidis* polypeptide include enzyme linked immunosorbent assays (ELISAs), Western blots, immunoprecipitations and immunofluorescence. Alternatively, *Alloiococcus otitidis*

polypeptides can be detected *in vivo* in a subject by introducing into the subject a labeled anti-*Alloicoccus otitidis* antibody. For example, the antibody can be labeled with a radioactive marker whose presence and location in a subject can be detected by standard imaging techniques.

5       The polynucleotides according to the invention may also be used in analytical DNA chips, which allow sequencing, the study of mutations and of the expression of genes, and which are currently of interest given their very small size and their high capacity in terms of number of analyses.

10       The principle of the operation of these chips is based on molecular probes, most often oligonucleotides, which are attached onto a miniaturized surface, generally of the order of a few square centimeters. During an analysis, a sample containing fragments of a target nucleic acid to be analyzed, for example DNA or RNA labeled, for example, after amplification, is deposited onto the DNA chip in which the support has been coated beforehand with probes. Bringing the labeled  
15       target sequences into contact with the probes leads to the formation, through hybridization, of a duplex according to the rule of pairing defined by J.D. Watson and F. Crick. After a washing step, analysis of the surface of the chip allows the effective hybridizations to be located by means of the signals emitted by the labels tagging the target. A hybridization fingerprint results from this analysis which, by appropriate  
20       computer processing, will make it possible to determine information such as the presence of specific fragments in the sample, the determination of sequences and the presence of mutations.

25       The chip consists of a multitude of molecular probes, precisely organized or arrayed on a solid support whose surface is miniaturized. It is at the center of a system where other elements (imaging system, microcomputer) allow the acquisition and interpretation of a hybridization fingerprint.

30       The hybridization supports are provided in the form of flat or porous surfaces (pierced with wells) composed of various materials. The choice of a support is determined by its physicochemical properties, or more precisely, by the relationship between the latter and the conditions under which the support will be placed during the synthesis or the attachment of the probes or during the use of the chip. It is therefore necessary, before considering the use of a particular support, to consider characteristics such as its stability to pH, its physical strength, its reactivity and its

chemical stability as well as its capacity to nonspecifically bind nucleic acids. Materials such as glass, silicon and polymers are commonly used. Their surface is, in a first step, called "functionalization", made reactive towards the groups which it is desired to attach thereon. After the functionalization, so-called spacer molecules are  
5 grafted onto the activated surface. Used as intermediates between the surface and the probe, these molecules of variable size render unimportant the surface properties of the supports, which often prove to be problematic for the synthesis or the attachment of the probes and for the hybridization.

Among the hybridization supports, there may be mentioned glass which is  
10 used, for example, in the method of *in situ* synthesis of oligonucleotides by photochemical addressing developed by the company Affymetrix (E.L. Sheldon, 1993), the glass surface being activated by silane. Genosensor Consortium (P. Mérel, 1994) also uses glass slides carrying wells 3 mm apart, this support being activated with epoxysilane.

15 The probes according to the invention may be synthesized directly *in situ* on the supports of the DNA chips. This *in situ* synthesis may be carried out by photochemical addressing (developed by the company Affymax (Amsterdam, Holland) and exploited industrially by its subsidiary Affymetrix (United States)) or based on the VLSIPS (very large scale immobilized polymer synthesis) technology  
20 (S.P.A. Fodor *et al.*, 1991) which is based on a method of photochemically directed combinatorial synthesis and the principle of which combines solid-phase chemistry, the use of photolabile protecting groups and photolithography.

The probes according to the invention may be attached to the DNA chips in various ways such as electrochemical addressing, automated addressing or the use  
25 of probe printers (T. Livache *et al.*, 1994; G. Yershov *et al.*, 1996; J. Derisi *et al.*, 1996, and S. Borman, 1996).

The revealing of the hybridization between the probes of the invention, deposited or synthesized *in situ* on the supports of the DNA chips, and the sample to be analyzed, may be determined, for example, by measurement of fluorescent  
30 signals, by radioactive counting or by electronic detection.

The use of fluorescent molecules such as fluorescein constitutes the most common method of labeling the samples. It allows direct or indirect revealing of the hybridization and allows the use of various fluorochromes.

Affymetrix currently provides an apparatus or a scanner designed to read its Gene Chip™ chips. It makes it possible to detect the hybridizations by scanning the surface of the chip in confocal microscopy (R.J. Lipshutz *et al.*, 1995).

The nucleotide sequences according to the invention are also used in DNA  
5 chips to carry out the analysis of the expression of the *Alloiococcus otitidis* genes. This analysis of the expression of *Alloiococcus otitidis* genes is based on the use of chips where probes of the invention, chosen for their specificity to characterize a given gene, are present (D.J. Lockhart *et al.*, 1996; D.D. Shoemaker *et al.*, 1996). For the methods of analysis of gene expression using the DNA chips, reference may,  
10 for example, be made to the methods described by D.J. Lockhart *et al.* (1996) and Sosnowsky *et al.* (1997) for the synthesis of probes *in situ* or for the addressing and the attachment of previously synthesized probes. The target sequences to be analyzed are labeled and in general fragmented into sequences of about 50 to 100 nucleotides before being hybridized onto the chip. After washing as described,  
15 for example, by D.J. Lockhart *et al.* (1996) and application of different electric fields (Sosnowsky *et al.*, 1997), the labeled compounds are detected and quantified, the hybridizations being carried out at least in duplicate. Comparative analyses of the signal intensities obtained with respect to the same probe for different samples and/or for different probes with the same sample, determine the differential  
20 expression of RNA or of DNA derived from the sample.

The nucleotide sequences according to the invention are, in addition, used in DNA chips where other nucleotide probes specific for other microorganisms are also present, and allow the carrying out of a serial test allowing rapid identification of the presence of a microorganism in a sample.

25 Accordingly, the subject of the invention is also the nucleotide sequences according to the invention, characterized in that they are immobilized on a support of a DNA chip.

The DNA chips, characterized in that they contain at least one nucleotide sequence according to the invention, immobilized on the support of the said chip,  
30 also form part of the invention.

The chips preferably contain several probes or nucleotide sequences of the invention of different length and/or corresponding to different genes so as to identify, with greater certainty, the specificity of the target sequences or the desired mutation

in the sample to be analyzed.

Accordingly, the analyses carried out by means of primers and/or probes according to the invention, immobilized on supports such as DNA chips, make it possible, for example, to identify, in samples, mutations linked to variations such as intraspecies variations. These variations may be correlated or associated with pathologies specific to the variant identified and make it possible to select the appropriate treatment.

The invention thus comprises a DNA chip according to the invention, characterized in that it contains, in addition, at least one nucleotide sequence of a microorganism different from *Alloiococcus otitidis*, immobilized on the support of the said chip; preferably, the different microorganism is chosen from an associated microorganism, a bacterium of the *Streptococcus* family, and a variant of the species *Alloiococcus otitidis*.

The principle of the DNA chip as explained above, is also used to produce protein "chips" on which the support has been coated with a polypeptide or an antibody according to the invention, or arrays thereof, in place of the DNA. These protein "chips" make it possible, for example, to analyze the biomolecular interactions (BIA) induced by the affinity capture of target analytes onto a support coated, for example, with proteins, by surface plasma resonance (SPR). Reference may be made, for example, to the techniques for coupling proteins onto a solid support which are described in EP 524 800 or to the methods describing the use of biosensor-type protein chips such as the BIAcore-type technique (Pharmacia) (Arlinghaus *et al.*, 1997, Krone *et al.*, 1997, Chatelier *et al.*, 1995). These polypeptides or antibodies according to the invention, capable of specifically binding antibodies or polypeptides derived from the sample to be analyzed, are thus used in protein chips for the detection and/or the identification of proteins in samples. The said protein chips may in particular be used for infectious diagnosis and preferably contain, per chip, several polypeptides and/or antibodies of the invention of different specificity, and/or polypeptides and/or antibodies capable of recognizing microorganisms different from *Alloiococcus otitidis*.

Accordingly, the subject of the present invention is also the polypeptides and the antibodies according to the invention, characterized in that they are immobilized on a support, in particular, on a protein chip.

The protein chips, characterized in that they contain at least one polypeptide or one antibody according to the invention immobilized on the support of the said chip, also form part of the invention.

The invention comprises, in addition, a protein chip according to the invention,  
5 characterized in that it contains, in addition, at least one polypeptide of a microorganism different from *Alloiococcus otitidis* or at least one antibody directed against a compound of a microorganism different from *Alloiococcus otitidis*, immobilized on the support of the chip.

The invention also relates to a kit or set for the detection and/or the  
10 identification of bacteria belonging to the species *Alloiococcus otitidis* or to an associated microorganism, or for the detection and/or the identification of a microorganism characterized in that it comprises a protein chip according to the invention.

The present invention also provides a method for the detection and/or the  
15 identification of bacteria belonging to the species *Alloiococcus otitidis* or to an associated microorganism in a biological sample, characterized in that it uses a nucleotide sequence according to the invention.

The invention also encompasses kits for detecting the presence of an  
*Alloiococcus otitidis* polypeptide in a biological sample. For example, the kit  
20 comprises reagents such as a labeled or labelable compound or agent capable of detecting *Alloiococcus otitidis* polypeptide or mRNA in a biological sample; means for determining the amount of *Alloiococcus otitidis* polypeptide in the sample; and means for comparing the amount of *Alloiococcus otitidis* polypeptide in the sample with a standard. The compound or agent are packaged in a suitable container. The kit  
25 further comprises instructions for using the kit to detect *Alloiococcus otitidis* mRNA or protein.

In certain embodiments, detection involves the use of a probe/primer in a polymerase chain reaction (PCR) (see, e.g. U.S. 4,683,195 and U.S. 4,683,202), such as anchor PCR or RACE PCR, or, alternatively, in a ligation chain reaction  
30 (LCR). This method includes the steps of collecting a sample of cells from a patient, isolating nucleic acid (e.g., genomic, mRNA or both) from the cells of the sample, contacting the nucleic acid sample with one or more primers which specifically hybridize to an *Alloiococcus otitidis* polynucleotide under conditions such that

hybridization and amplification of the *Alloiococcus otitidis*-polynucleotide (if present) occurs, and detecting the presence or absence of an amplification product, or detecting the size of the amplification product and comparing the length to a control sample.

5

#### I. TRANSGENIC ANIMALS

It is contemplated that in some instances the genome of a transgenic animal of the present invention will have been altered through the stable introduction of one or more of the *Alloiococcus otitidis* polynucleotide compositions described herein, either native, synthetically modified or mutated. As described herein, a "transgenic animal" refers to any animal, preferably a non-human mammal (*e.g.* mouse, rat, rabbit, squirrel, hamster, rabbits, guinea pigs, pigs, micro-pigs, baboons, squirrel monkeys and chimpanzees, *etc.*), bird or an amphibian, in which one or more cells contain a heterologous nucleic acid sequence introduced by way of human intervention, such as by transgenic techniques well known in the art. The nucleic acid is introduced into the cell, directly or indirectly, by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. The term genetic manipulation does not include classical crossbreeding, or *in vitro* fertilization, but rather is directed to the introduction of a recombinant DNA molecule. This molecule may be integrated within a chromosome, or it may be extrachromosomally replicating DNA.

The host cells of the invention are also used to produce non-human transgenic animals. The non-human transgenic animals are used in screening assays designed to identify infections or compounds, *e.g.*, drugs, pharmaceuticals, *etc.*, which are capable of ameliorating *Alloiococcus otitidis* symptoms or infections. For example, in one embodiment, a host cell of the invention is a fertilized oocyte or an embryonic stem cell into which an *Alloiococcus otitidis* polypeptide-coding sequence has been introduced. Such host cells are then used to create non-human transgenic animals in which exogenous *Alloiococcus otitidis* gene sequences have been introduced into their genome or homologous recombinant animals in which endogenous *Alloiococcus otitidis* gene sequences have been altered. Such animals are useful for studying the effects of an *Alloiococcus otitidis* polypeptide and for

identifying and/or evaluating modulators of *Alloiococcus otitidis* polypeptide infectivity.

A transgenic animal of the invention is created by introducing an *Alloiococcus otitidis* polypeptide-encoding nucleic acid sequence into the male pronuclei of a fertilized oocyte, *e.g.*, by microinjection, retroviral infection, and allowing the oocyte to develop in a pseudopregnant female foster animal. The human *Alloiococcus otitidis* cDNA sequence of one or more of SEQ ID NO:1 through SEQ ID NO: 4035 can be introduced as a transgene into the genome of a non-human animal.

Moreover, a non-*Alloiococcus otitidis* homologue of the *Alloiococcus otitidis* gene can be isolated based on hybridization to the *Alloiococcus otitidis* cDNA (described above) and used as a transgene. Intronic sequences and polyadenylation signals can also be included in the transgene to increase the efficiency of expression of the transgene. A tissue-specific regulatory sequence(s) can be operably linked to the *Alloiococcus otitidis* transgene to direct expression of an *Alloiococcus otitidis* polypeptide to particular cells. Methods for generating transgenic animals *via* embryo manipulation and microinjection, particularly animals such as mice, have become conventional in the art and are described, for example, in U.S. 4,736,866 and 4,870,009, U.S. 4,873,191 and in Hogan, 1986. Similar methods are used for production of other transgenic animals. A transgenic founder animal can be identified based upon the presence of the *Alloiococcus otitidis* transgene in its genome and/or expression of *Alloiococcus otitidis* mRNA in tissues or cells of the animals. A transgenic founder animal can then be used to breed additional animals carrying the transgene. Moreover, transgenic animals carrying a transgene encoding an *Alloiococcus otitidis* polypeptide can further be bred to other transgenic animals carrying other transgenes.

In another embodiment, transgenic non-human animals can be produced which contain selected systems that allow for regulated expression of the transgene. One example of such a system is the cre/loxP recombinase system of bacteriophage P $\lambda$ . For a description of the cre/loxP recombinase system, *see, e.g.*, Lakso *et al.*, 1992. Another example of a recombinase system is the FLP recombinase system of *Saccharomyces cerevisiae* (O'Gon-nan *et al.*, 1991). If a cre/loxP recombinase system is used to regulate expression of the transgene, animals containing transgenes encoding both the Cre recombinase and a selected protein are required.



Such animals can be provided through the construction of "double" transgenic animals, *e.g.*, by mating two transgenic animals, one containing a transgene encoding a selected protein and the other containing a transgene encoding a recombinase.

5 Clones of the non-human transgenic animals described herein can also be produced according to the methods described in Wilmut *et al.*, 1997, and PCT International Publication Nos. WO 97/07668 and WO 97/07669. In brief, a cell, *e.g.*, a somatic cell, from the transgenic animal can be isolated and induced to exit the growth cycle and enter G<sub>0</sub> phase. The quiescent cell can then be fused, *e.g.*, through  
10 the use of electrical pulses, to an enucleated oocyte from an animal of the same species from which the quiescent cell is isolated. The reconstructed oocyte is then cultured such that it develops to morula or blastocyst and then transferred to pseudopregnant female foster animal. The offspring borne of this female foster animal will be a clone of the animal from which the cell, *e.g.*, the somatic cell, is  
15 isolated.

All patents and publications cited herein are hereby incorporated by reference.

The following examples are carried out using standard techniques, which are well known and routine to those of skill in the art, except where otherwise described  
20 in detail. The following examples are presented for illustrative purposes, and should not be construed in any way limiting the scope of this invention.

#### EXAMPLE 1

##### CONFIRMATION OF THE IDENTITY OF THE *ALLOIOCOCCUS OTITIDIS* 1104-92 ISOLATE

25 The *Alloiococcus otitidis* isolate 1104-92 was obtained from Dr. Richard Facklam of the Centers for Disease Control in Atlanta. It was isolated from the middle ear fluid of a child in the Atlanta, Georgia area. It was confirmed to be *A. otitidis* by comparing it to the type strain, ATCC51267, obtained from the American Type  
30 Culture Collection [Aguirre, 1992 #1]. Both the 1104-92 and type strain are characterized as Gram positive cocci. Both grow on Columbia agar supplemented with 5% yeast extract, 0.5% polysorbate 80 (Tween 80), and 0.7% phosphatidyl choline when incubated at 37°C. On this medium, both strains form slow growing

small white colonies that require nearly two days to be easily observed with the naked eye. Both are sensitive to lysis by hen egg white lysozyme and *Streptococcus globisporus* mutanolysin. Both grow in the presence of 2% sodium azide. Both are killed by incubation at 55°C for 30 minutes. Finally, to further confirm that the 1104-92 was a strain of *A. otitidis*, it was subject to polymerase chain reaction (PCR) identification based on its 16s rRNA gene. This was done using two of the primers specified by Aguirre and Collins [Aguirre, 1992 #2]. The antisense primer used was 5'-ATCTTCCTGCTTGCAGGAAGAGG-3' and the sense primer was 3'-CGCTTCATCTCTGAAGCTAGC-5'. Thus by multiple criteria, the 1104-92 strain was confirmed to be an isolate of *A. otitidis*.

#### EXAMPLE 2

##### STORAGE, GROWTH, AND HARVEST OF *ALLOIOCOCCUS OTITIDIS* 1104-92 FOR ISOLATION OF DNA

The *A. otitidis* isolate 1104-92 was stored at -70°C in Todd-Hewlett broth containing 40% glycerol. A small portion of the frozen stock was streaked onto the agar medium described in Example 1 and incubated at 37°C for two days. The growth from the plate was swabbed into a 17 × 100 cm tube containing 6 ml of a serum-free broth medium. This broth medium was prepared with 30 g Todd-Hewlett medium, 5 g yeast extract, 10 ml polysorbate 80 (Tween 80), and 1 liter distilled water. This medium was sterilized by autoclaving for 35 minutes. The bacteria were incubated aerobically without shaking in an aerobic incubator at 37°C for two days. The tube containing the growing bacteria was then shaken to resuspend the bacteria and added to a liter of the same medium in a Fernbach flask. This flask, in turn, was incubated aerobically for three days without shaking. The bacteria were harvested by first swirling the flask to suspend the bacteria and then low speed centrifugation at about 5,000 × g for 30 minutes. The pellet of bacteria was washed by resuspending it in 10 to 15 mL of phosphate buffered saline (PBS), and centrifuging the suspension at about 8,000 × g for 20 minutes. The pellet of bacteria was retained and stored frozen at -20°C. The yield of wet bacterial pellet was typically about 1 g per liter of broth.

## EXAMPLE 3

PREPARATION OF ALLOIOCOCCUS OTITIDIS GENOMIC DNA

To prepare genomic DNA, 0.95 g frozen pellet of bacteria was defrosted and  
5 suspended in 10 mL of PBS containing 1 mM MgCl<sub>2</sub>. The bacteria were killed by  
incubating the suspension at 55°C for 20 minutes. The suspension was allowed to  
cool before adding 25 µl of a 10 mg/mL stock of hen egg white lysozyme and 50 µl of  
a 25,000 unit/mL stock of *Streptococcus globisporus* mutanolysin to the suspension.  
It was then incubated for one hour at 37°C. Then 50 µl of a 10 mg/mL stock of  
10 RNase was added and the suspension incubated an additional hour at 37°C. After  
these incubations, sodium dodecylsulfate (SDS) was added to a final concentration  
of 0.3% (0.3 mL of a 10% stock). This was followed by the addition of 0.3 mL of a 1  
mg/mL stock of proteinase K. The suspension was then incubated for two hours at  
37°C. After this time, an equal volume of water saturated phenol/chloroform/isopropyl  
15 (25:24:1) was added to the digested suspension and gently mixed. The upper  
aqueous layer was retained after a low speed centrifugation and 2.5 volumes of  
ethanol were added and the tube gently inverted to mix. The DNA was then spooled  
out on a glass rod and allowed to air dry.

The DNA at this stage still contained obvious impurities and needed further  
20 purification. The DNA dried on the glass rod was soaked in 70% ethanol to remove  
excess phenol and air-dried once again. It was then suspended in 2 ml of Tris-EDTA  
buffer to which 2 µl of RNase cocktail was added and incubated at room temperature  
for 75 minutes. Then 100 µl of protease, 100 µl SDS and 40 µl of 100 mM CaCl<sub>2</sub>  
were added and the suspension incubated for 3.5 hours. An equal volume of  
25 chloroform was added, gently mixed, then centrifuged at a low speed. The aqueous  
layer was collected and re-extracted with the phenol, chloroform, isopropyl alcohol  
reagent. In turn, the aqueous layer was extracted with chloroform. At this point, 3 M  
sodium acetate was added to the aqueous phase collected from the last extraction  
and then 3.75 ml of ethanol was added and gently mixed. The DNA was spooled out,  
30 soaked in 70% ethanol and allowed to air-dry. The DNA was finally suspended in 2  
ml of Tris-EDTA buffer. Based on absorption at 260 nm, the final yield of DNA was  
482 µg of DNA. The DNA was confirmed to be that of *A. otitidis* by the PCR method  
described in example 1. This DNA was submitted for sequencing.

## EXAMPLE 4

CLONING AND SEQUENCING ALLOIOCOCCUS OTTIDIS GENOME

5 This invention provides nucleotide sequences of the genome of *Alloiococcus otitidis* which thus comprises a DNA sequence library of *Alloiococcus otitidis* genomic DNA. The detailed description that follows provides nucleotide sequences of *Alloiococcus otitidis*, and also describes how the sequences were obtained and how ORFs (Open Reading Frames) and protein-coding sequences can be identified.

10 To construct a library, genomic DNA was hydrodynamically sheared in an HPLC and then separated on a standard 1% agarose gel. A fraction corresponding to 3000-3500 bp in length was excised from the gel and purified by the GeneClean procedure (BIO101, Inc.).

The purified DNA fragments were then blunt-ended using T4 DNA  
15 polymerase. The blunt-ended DNA was then ligated to unique BstX1-linker adapters. These linkers are complimentary to the pGTC vector, while the overhang is not self-complimentary. Therefore, the linkers will not concatemerize nor will the cut-vector religate itself easily. The linker-adapted inserts were separated from the unincorporated linkers on a 1% agarose gel and again purified using GeneClean.  
20 The linker-adapted inserts were then ligated to BstX1-cut vector to construct "shotgun" subclone libraries.

Only major modifications to the protocols are highlighted. Briefly, the library was transformed into DH10B competent cells (Gibco/BRL, DH5a transformation protocol). Transformed cells were detected by plating onto antibiotic plates  
25 containing ampicillin. The plates were incubated overnight at 37° C. Transformant clones were then selected for sequencing. The cultures were grown overnight at 37°C. DNA was purified using a silica bead DNA preparation (Egelstein, 1996) method. In this manner, 25 mg of DNA was obtained per clone.

These purified DNA samples were then sequenced using ABI dye-terminator  
30 chemistry. All subsequent steps were based on sequencing by automated DNA sequencing methods. The ABI dye terminator sequence reads were run on MegaBace™ 10000 (Amersham) machines and the data transferred to UNIX based computers. Base calls and quality scores were determined using the PHRED

software program (Ewing et al., 1998, Genome Res. 8: 175-185; Ewing and Green, 1998, Genome Res. 8:685-734). Reads were assembled using PHRAP (P. Green, Abstracts of DOE Human Genome Program Contractor-Grantee Workshop V, Jan. 1996, p 157) with default program parameters and quality scores.

5 To identify *Alloiococcus otitidis* genome encoded polypeptides, the complete genomic sequence of *Alloiococcus otitidis* was analyzed essentially as follows: First, all possible stop-to-stop open reading frames (ORFs)  $\geq 222$  nucleotides in all three reading frames were translated into amino acid sequences.

Second, the identified ORFs were analyzed for homology to known protein  
10 sequences. Third, the coding potential of non-homologous sequences were evaluated with the GENEMARK<sup>TM</sup> software program (Borodovsky and McIninch, 1993, Comp. Chem. 17:123). The results of these analysis are set forth in tables 2-16.

#### 15 EXAMPLE 5

##### IDENTIFICATION OF SPECIFIC GENES IN ALLOIOCOCCUS OTITIDIS

*Alloiococcus otitidis* homologs of the genes listed in Table 4 were identified as follows:

20 Protein sequences of interest ("query sequences", Table 4) were extracted from Genbank from one or more species; query species included but were not limited to *Staphylococcus aureus*, *Streptococcus pneumoniae*, *Streptococcus pyogenes*, *Lactococcus lactis*, *Escherichia coli*, and *Bacillus subtilis*. These queries were compared to the *Alloiococcus otitidis* sequence by several methods in order to  
25 determine which *Alloiococcus* sequence was the ortholog for the query gene.

First, the query sequences were compared to the translated *Alloiococcus otitidis* ORF set using BLASTP. The ORF set was generated as described in Vaccines patent, except that for each ORF that had multiple potential start codons, only the longest ORF was used. The top 10 *Alloiococcus otitidis* hits for each query  
30 were saved, without regard to score.

These *Alloiococcus otitidis* hits were then compared to NR, the nonredundant Genpept database, using BLASTP. An *Alloiococcus otitidis* ORF was considered the ortholog of a query sequence if the genes were reciprocal best hits in *Alloiococcus*

*otitidis* and the query genome. This analysis is also summarized in Table 4 (excel file AOT\_PATENT\_FILE.xls, Sheet TopHitsAndClustalKey). Specific numerical cutoffs were not used; however all top hits had Expect values of less than  $3 \times 10^{-28}$ .

Several query sequences had more than one high-scoring hit in *Alloiococcus*  
5 *otitidis*. In most cases, however, only the first, best hit to the original query sequence had that query sequence as its reciprocal best hit. For example, the *Streptococcus pyogenes* query sequence GyrA (alpha subunit of DNA gyrase) has two high-scoring hits in *Alloiococcus otitidis*. These were distinguished by the reciprocal blast analysis; the first, ORF\_505 (60% identity, Expect = 0) is the GyrA homolog and the  
10 second, ORF\_1907 (38% identity, Expect =  $1 \times 10^{-154}$ ) is the homolog of the query sequence GrlA or ParC (topoisomerase IV, A subunit). Other examples of closely related proteins include the B subunits of DNA gyrase (GyrB) and Topoisomerase IV (GrlB or ParE); and YphC and Era, both of which are putative GTP binding proteins of unknown function. These *Alloiococcus otitidis* ORFS were assigned based on  
15 their top hit in Genpept.

In two cases the multiple high-scoring hits in *Alloiococcus otitidis* were the result of gene duplication. In the case of MurA (UDP-N-acetylglucosamine enolpyruvyl transferase) two separate *Alloiococcus otitidis* ORFS were determined to be the desired orthologs, because both had MurA (or MurZ, alternate notation) as  
20 their best hit in Genpept. Likewise, there are two FolC (folypolyglutamate synthase) homologs in *Alloiococcus otitidis*. It is known that other bacteria, particularly Gram-positive bacteria, may carry two homologs of each of these genes.

As a further step in verification of gene assignments, the *Alloiococcus otitidis* ORFS identified as orthologs of the query genes by the analysis above were then  
25 compared to an internal copy of the COGS database (Tatusov RL, Natale DA, Garkavtsev IV, Tatusova TA, Shankavaram UT, Rao BS, Kiryutin B, Galperin MY, Fedorova ND, Koonin EV, 2001, Nucleic Acids Res 2001 Jan 1;29(1):22-8. The COG database: new developments in phylogenetic classification of proteins from complete genomes) using BLASTP. The COGS database is a curated set of proteins  
30 from a set of finished bacterial genomes, which have been grouped into specific protein families on the basis of protein similarity. In all cases, the *Alloiococcus otitidis* ORF was most closely related to the COGS family of the initial query protein, if that protein had been assigned to a COGS family. Examples of proteins for which

there is no COGS family defined (in our local version of the database) include SrtA (sortase) and MvaK1 (phosphomevalonate kinase).

As a final confirmation, all query proteins were compared to the complete *Alloiococcus otitidis* nucleotide sequence using TBLASTN, in order to determine if there were additional and/or better hits that had not been predicted as ORFS. In all cases, the same sequence was identified as the best hit by TBLASTN and by BLASTP.

For one query sequence, sortase, the *Alloiococcus otitidis* ORF that was the top hit (Expect = 0.42) by the initial BLASTP or TBLASTN using the *Staphylococcus aureus* sortase sequence as query was found by additional analysis (reciprocal blast) to be a putative ABC-transport protein. The true sortase homolog in *Alloiococcus otitidis* was identified by construction of a Hidden Markov Model based on a multiple alignment of 72 known and putative sortase proteins that had been identified previously using similar computational methods. The model was constructed using "hmmbuild" and the *Alloiococcus otitidis* ORF set was searched using "hmmsearch", both of the hmmer package (S.R. Eddy. Profile hidden Markov models. *Bioinformatics* 14:755-763, 1998). The assignment of ORF\_876 as sortase was then confirmed by reciprocal blast as described above and in Table 2. ORF\_876 was also found to be the top hit in *Alloiococcus otitidis* when the *Bacillus subtilis* putative sortase (YhcS) was used as the query sequence in a BLASTP search. The *Bacillus halodurans* BH3596 *Bacillus subtilis* YhcS and proteins that are the top hits for RF\_876 have recently been placed into a COGS group of sortases, further confirming the identity of ORF\_876 as the *Alloiococcus otitidis* sortase.

TABLE 4

ORF NO.	DNA SEQ ID NO.r	ORF Start	Protein Start	ORF Stop	Protein SEQ ID No.	Gene
46b	Seq. ID No. 1	26225	26153	25800	Seq. ID No. 2	ACPS
48c	Seq. ID No. 3	29105	29090	27696	Seq. ID No. 4	murF
57b	Seq. ID No. 5	33738	33732	32455	Seq. ID No. 6	murA-2
65c	Seq. ID No. 7	37245	37242	36634	Seq. ID No. 8	rpoE
172	Seq. ID No. 9	88726	88726	87785	Seq. ID No. 10	rpoA
228d	Seq. ID No. 11	111563	111542	107883	Seq. ID No. 12	rpoC
236c	Seq. ID No. 13	115224	115221	111643	Seq. ID No. 14	rpoB
495c	Seq. ID No. 15	247355	247331	245949	Seq. ID No. 16	dnaB/C
505b	Seq. ID No. 17	254277	254268	251794	Seq. ID No. 18	gyrA
509	Seq. ID No. 19	256252	256246	254297	Seq. ID No. 20	gyrB
515c	Seq. ID No. 21	259131	259116	257914	Seq. ID No. 22	dnaN
528b	Seq. ID No. 23	263837	263861	265153	Seq. ID No. 24	folC-2
851b	Seq. ID No. 25	440982	441072	442634	Seq. ID No. 26	murE
876	Seq. ID No. 27	453874	453898	454509	Seq. ID No. 28	srtA
956	Seq. ID No. 29	500019	500019	501308	Seq. ID No. 30	folC-1
959	Seq. ID No. 31	501978	501993	502364	Seq. ID No. 32	folB
961c	Seq. ID No. 33	502392	502413	502943	Seq. ID No. 34	folK
1183b	Seq. ID No. 35	626391	626430	627632	Seq. ID No. 36	mvaS
1184	Seq. ID No. 37	629315	629285	627993	Seq. ID No. 38	mvaA
1263	Seq. ID No. 39	675596	675608	676525	Seq. ID No. 40	murB
1273	Seq. ID No. 41	685392	685377	684289	Seq. ID No. 42	mvaK2
1275b	Seq. ID No. 43	686415	686403	685393	Seq. ID No. 44	mvaD
1277	Seq. ID No. 45	687376	687349	686396	Seq. ID No. 46	mvaK1
1279	Seq. ID No. 47	687461	687506	688435	Seq. ID No. 48	coaA
1284	Seq. ID No. 49	691675	691681	692520	Seq. ID No. 50	nadE
1511	Seq. ID No. 51	815078	815084	815920	Seq. ID No. 52	murI
1811b	Seq. ID No. 53	985498	985504	986454	Seq. ID No. 54	folP
1863b	Seq. ID No. 55	1019023	1019050	1019583	Seq. ID No. 56	folA
1902	Seq. ID No. 57	1040639	1040645	1042606	Seq. ID No. 58	GrlB
1907	Seq. ID No. 59	1042729	1042732	1045191	Seq. ID No. 60	grlA



Table 4  
(Cont'd.)

ORF NO.	DNA SEQ ID NO.r	ORF Start	Protein Start	ORF Stop	Protein SEQ ID No.	Gene
1990c	Seq. ID No. 61	1098801	1098798	1097689	Seq. ID No. 62	rpoD
1992b	Seq. ID No. 63	1100670	1100670	1098817	Seq. ID No. 64	dnaG
2003	Seq. ID No. 65	1109198	1109144	1108212	Seq. ID No. 66	era
2016h	Seq. ID No. 67	1115435	1115390	1113879	Seq. ID No. 68	norA
2133	Seq. ID No. 69	1179995	1179938	1175604	Seq. ID No. 70	polC
2181b	Seq. ID No. 71	1203606	1203588	1202281	Seq. ID No. 72	obg
2204	Seq. ID No. 73	1216828	1216804	1215491	Seq. ID No. 74	ypnC
2240c	Seq. ID No. 75	1236616	1236607	1233293	Seq. ID No. 76	dnaE
2284	Seq. ID No. 77	1261069	1261063	1259858	Seq. ID No. 78	coaBC
2328	Seq. ID No. 79	1286689	1286668	1285637	Seq. ID No. 80	holA
2333	Seq. ID No. 81	1290847	1290847	1290371	Seq. ID No. 82	coaD
2485	Seq. ID No. 83	1374427	1374400	1373168	Seq. ID No. 84	ftsZ
2489	Seq. ID No. 85	1375804	1375792	1374428	Seq. ID No. 86	ftsA
2492b	Seq. ID No. 87	1378075	1378060	1376897	Seq. ID No. 88	murG
2494	Seq. ID No. 89	1379477	1379453	1378050	Seq. ID No. 90	murD
2514	Seq. ID No. 91	1390141	1390135	1389491	Seq. ID No. 92	nadD
2596	Seq. ID No. 93	1437374	1437374	1436709	Seq. ID No. 94	coaE
2602	Seq. ID No. 95	1442399	1442396	1441065	Seq. ID No. 96	murC
2645	Seq. ID No. 97	1467800	1467782	1466751	Seq. ID No. 9	fmhB
2875	Seq. ID No. 99	1605944	1605923	1603701	Seq. ID No. 100	pcrA
2918	Seq. ID No. 101	1631092	1631089	1629779	Seq. ID No. 102	murA-1
3001	Seq. ID No. 103	1680254	1680221	1679229	Seq. ID No. 104	holB
3012	Seq. ID No. 105	1684114	1684102	1682330	Seq. ID No. 106	dnaX

5

## EXAMPLE 6

IDENTIFICATION OF THE GENE ENCODING COENZYME A (COA) IN ALLOIOCOCCUS OTTIDIS

10

Pantothenate kinase (PanK, CoaA) encoded by the *coaA* gene catalyzes the initial step in Coenzyme A (CoA) biosynthesis. CoA is an essential co-factor in a number of metabolic pathways in bacteria and mammals. Short-chain thioesters such as acetyl-CoA and succinyl-CoA are essential intermediates in carbon metabolism. CoA-thioesters of long chain fatty acids feed into  $\beta$ -oxidation and are also the source of fatty acids for phospholipids. In addition, CoA and its thioesters play important roles in the regulation of several enzymes in intermediary metabolism, including pyruvate dehydrogenase and phosphoenolpyruvate carboxylase. Finally,

synthesis of holo acyl carrier protein (ACP) is dependent on CoA for the 4'-phosphopantetheine moiety linked to ACP. ACP is essential for fatty acid biosynthesis. The two major acyl-carrier groups in cells: CoA and ACP, are derived from pantothenate. Pantothenate can be obtained exogenously through uptake via a permease, the product of the *panF* gene. Alternately, pantothenate is the product of condensation of pantoate and  $\beta$ -alanine via pantothenate synthetase, the product of the *panC* gene. The initial step in CoA biosynthesis is the phosphorylation of pantothenate by pantothenate kinase (PanK, CoaA).

The *coaA* gene was originally identified by Dunn and Snell in *S. typhimurium* as a temperature sensitive allele. Similarly, a temperature sensitive allele of *coaA* was reported for *E. coli* in 1987. CoaA was found to be essential in *E. coli* in a recent genetic footprinting analysis. In the temperature sensitive strains, accumulation of phosphorylated CoA intermediates rapidly ceased following shift to the non-permissive temperature. CoaA was shown to be a homo-dimer of 35 kDa subunits that bound ATP cooperatively. ATP is bound first in a sequential mechanism of action; CoA has been shown to be a potent inhibitor of the reaction and competitively competes for binding with ATP. Therefore CoaA is under feedback regulation and is the major regulatory step in CoA biosynthesis.

Lysine 101 in bacterial pantothenate kinase (CoaA) was found to be essential for both ATP and CoA binding. This supports kinetic data that CoA is a competitive inhibitor of ATP binding to CoaA and that both substrates bind to the same site.

Homologues of *E. coli* CoaA have been identified in *B. subtilis*, *S. pyogenes*, *M. tuberculosis*, *H. influenzae* and *V. cholerae*. Homologues have not been identified in either the *S. cerevisiae* genome or in a mammalian expressed sequence tag database. Calder *et al.* identified a homologue, through functional complementation of an *E. coli coaA* ts mutant, in *A. nidulans*. Homologue of this gene identified in *Alloiococcus otitidis* as described in Example 5 (Seq. ID No 47. The protein encoded by the gene is set forth in Seq. ID No. 48.

The *A. nidulans* gene was then used to identify a yeast homologue. The bacterial and *Aspergillus* enzymes were found to be 16% identical and 32% similar. Although this level of similarity is quite weak the essential lysine residue involved in nucleotide binding appears to be conserved; however, the sequence surrounding the lysine residue were not conserved and further study will be required to validate this

finding. The most striking difference between the eukaryotic and prokaryotic enzymes is found in the sensitivity of each to competitive inhibition by CoA and acetyl-CoA. The yeast enzyme was most sensitive to acetyl-CoA and less sensitive to CoA, whereas the converse was true for the bacterial enzyme. Later studies  
5 demonstrated that mammalian pantothenate kinase is activated by CoA and inhibited by acetyl-CoA.

#### Nucleotide binding

Binding of ATP to CoaA is directly demonstrated by equilibrium dialysis  
10 employing the non-hydrolyzable ATP analogue ATP $\gamma$ S. The  $K_d$  measured for ATP binding is reported to be 2.1  $\mu$ M.

#### CoA binding

Binding of CoA to CoaA is directly demonstrated by equilibrium dialysis and  
15 the  $K_d$  is reported to be 6.7  $\mu$ M.

#### Pantothenate kinase activity

Specific kinase activity of CoaA is demonstrated using D-[1- $^{14}$ C]pantothenate and capturing 4'-phospho[1- $^{14}$ C]pantothenate on DE81 filters. Using this assay the  
20 following kinetic values were derived: specific activity – 470 $\pm$  200 nmol/min/mg; pantothenate  $K_m$  – 36  $\mu$ M;  $K_m$  ATP – 136  $\mu$ M.

#### Suitability of target for anti-infective development

Coenzyme A biosynthesis is essential for bacterial viability. CoaA catalyzes the  
25 first step of biosynthesis of CoA and appears to be the point of regulation for the pathway. The essentiality of CoaA is demonstrated through the construction of temperature sensitive alleles in *coaA*. Although the yeast enzyme is found to functionally complement the bacterial temperature sensitive allele, sequence and kinetic differences suggest the possibility of identifying inhibitors of the bacterial  
30 enzyme with high selectivity. As CoaA is essential and conserved in gram-negative and gram-positive pathogens, such inhibitors will have broad-spectrum utility.

**Suitable assays for measuring CoaA function**

CoaA is purified by standard methods using widely available molecular tags following expression at high level from *E. coli*. Pantothenate kinase activity is measured as follows: CoaA and D-[1-<sup>14</sup>C]pantothenate is incubated in a buffer consisting of 100 mM Tris (pH 7.4), 2.5 mM MgCl<sub>2</sub>, 2.5 mM ATP for 5-60 minutes at 37°C. Product, 4'-phospho[1-<sup>14</sup>C] pantothenate, is monitored through retention of labeled material on DE81 filters. This assay is amenable to high-throughput screening using high-density well-filter plates.

**EXAMPLE 7****IDENTIFICATION OF THE GENE ENCODING COABC (DFP) IN *ALLOIOCOCCUS OTTIDIS***

The *E. coli dfp* gene, which encodes the previously designated Dfp protein, was originally identified as encoding an enzyme required for CoA biosynthesis. The gene, coding for the protein of interest, was renamed *coaBC* to reflect the enzyme function in CoA biosynthesis. CoA is an essential co-factor in a number of metabolic pathways in bacteria and mammals. Short-chain thioesters such as acetyl-CoA and succinyl-CoA are essential intermediates in carbon metabolism.

CoaBC carries out the second and third steps of coenzyme A biosynthesis: the conjugation of 4'-phosphopantetheate with cysteine by the CoaB (PPCS : 4'-phosphopantethenoyl cysteine synthase) activity followed by the conversion to 4'-phosphopantetheine by the CoaC (PPCDC: 4'-phosphopantenoylcysteine decarboxylase) activity. Homologue of this gene identified in *Alloiococcus otitidis* as described in Example 5 (Seq. ID No 77). The protein encoded by the gene is set forth in Seq. ID No. 78.

**Enzyme activity of CoaBC (Dfp):**

Initially it was demonstrated that Dfp enzyme catalyzing oxidative decarboxylation of (R)-4'-phospho-N-pantothenoylcysteine (PPC) to form 4'-phosphopantetheine (PP) – the third step in CoA biosynthesis from pantothenate. The  $K_m$  for this reaction is 800  $\mu$ M for PPC.

Subsequently, it was established that Dfp is a bifunctional enzyme, catalyzing the second step of CoA biosynthesis, coupling of 4'-phosphopantothenate with

cysteine to form PPC, as well. This reaction is a two-step process and requires CTP for initial 4'-phosphopantothenate activation. Second step couples cysteine to the phosphopantothenate moiety with a release of CMP. Estimated  $K_M$ 's are 300  $\mu$ M for 4'-phosphopantothenate and CTP, and 250  $\mu$ M for cysteine.

5

#### **CoaBC as target for antibacterial development.**

Coenzyme A (CoA) plays a vital role in the metabolism of living cells. According to a recent report, 4% of all enzymes in the cell require CoA, its thioesters or 4'-phosphopantetheine. Recent genetic footprinting experiments on *E. coli* and direct gene knockout have established that this *coaBC* is essential for bacterial growth. Homologs of *coaBC* have been identified in a number of gram-positive and gram-negative organisms, which suggested the possibility of developing a broad-spectrum antibacterials from *coaBC* inhibitors. Considering the bifunctional nature of CoaBC, it is feasible to identify inhibitors that will inhibit both enzymic functions, thus arresting two steps in the CoA pathway. Another important factor in favor of selecting CoaBC as a target for antibacterials is low homology of the bacterial enzyme to eukaryotic counterparts. In most of the higher organisms including humans, two separate enzymes carry out these functions. Moreover, mammalian (R)-4'-phospho-N-pantothenoylecysteine decarboxylase is a pyruvate-dependent enzyme, while CoaBC requires flavine mononucleotide for its function.

10  
15  
20

#### **Assays for measuring CoaBC activity.**

PPC synthetase activity is be monitored by detecting the released pyrophosphate. This is achieved by converting pyrophosphate to inorganic phosphate with pyrophosphatase and detection by the Malachite Green assay, or by the MESG assay spectrophotometrically. CoaBC (2  $\mu$ g) is incubated in the reaction buffer containing 10 mM DTT, 2 mM  $MgCl_2$ , 50 mM Tris-HCl, pH 8, 300  $\mu$ M 4'-phosphopantothenate, 3.5 mM CTP, 5  $\mu$ g pyrophosphatase. The reaction is started by addition of appropriate amount (10-500  $\mu$ M final) of cysteine. The reaction is stopped at different time points by addition of equal volume of 5M  $H_2SO_4$ . The

25  
30

amount of inorganic phosphate released will be determined according to the one of described techniques.

PPC synthetase activity is also monitored by detecting the release of carbon dioxide from  $^{14}\text{C}$ -labeled cysteine. CoaBC (2  $\mu\text{g}$ ) is incubated in the reaction buffer containing 10 mM DTT, 2 mM  $\text{MgCl}_2$ , 50 mM Tris-HCl, pH 8, 2.5  $\mu\text{M}$  4'-phosphopantothenate, 3.5 mM CTP. The reaction is started by addition of appropriate amount (30 mM, final concentration) of  $^{14}\text{C}$ -labeled cysteine. The reaction is stopped at different time points by addition of equal volume of 5M  $\text{H}_2\text{SO}_4$ . Amount of released  $^{14}\text{C}$ -labeled  $\text{CO}_2$  is determined according to published technique.

### Example 8

#### Identification of the gene encoding phosphopantetheine adenylyltransferase (CoaD) in *Alloicoccus otitidis*

Phosphopantetheine adenylyltransferase, (PPAT, CoaD, KdtB) catalyzes the penultimate step in Coenzyme A (CoA) biosynthesis. The fourth step in CoA biosynthesis is the addition of AMP to 4'-phosphopantetheine by phosphopantetheine adenylyltransferase (CoaD) to form 3' dephospho-CoA (dPCoA).

The *coaD* gene was first identified in *E. coli* by Geerlof *et al.* CoaD is essential for viability in *E. coli* and *S. aureus*. The enzyme has a mass of 18 kDa and was determined to be a hexamer through cross-linking studies. Crystallography confirmed the oligomeric state of the enzyme. Moreover, co-crystallography of CoaD with dPCoA has also been carried out mapping the binding pocket for the major product of the reaction. Interestingly, in mammals PPAT has been shown to be in a complex with dephospho Coenzyme A kinase (dPCoA kinase, DPCK). This enzyme, purified from pig liver, is referred to as CoA Synthase. The yeast PPAT is associated with a protein complex that is in excess of 375 kDa and composed of six proteins. There is no detectable homology between the bacterial PPAT (CoaD) and the recently identified human PPAT, the activity of which is contained in a bifunctional PPAT/DPCK enzyme. Homologues of *E. coli* CoaD have been identified in *P. aeruginosa*, *S. pneumoniae*, *S. aureus*, *H. influenzae*, *H. pylori*, *B. anthracis* and *M. tuberculosis*. Homologue of this gene identified in *Alloicoccus otitidis* as described in

Example 5 (Seq. ID No 81). The protein encoded by the gene is set forth in Seq. ID No. 82.

#### **Enzyme activity**

5 CoaD (PPAT) carries out the reversible transfer of AMP to 4'-phosphopantetheine, forming dephosphocoenzyme A and releasing PP<sub>i</sub>. The reverse reaction was demonstrated by Geerlof *et al.* using a coupled assay to tie ATP production to NADP reduction, which is monitored at 340 nm. The following kinetic constants were calculated:  $k_{cat} = 3.3 \pm 0.1$  /sec;  $K_{m(dPCoA)} = 7.0 \pm 1.4$   $\mu$ M;  
10  $K_{m(PP_i)} = 0.22 \pm 0.04$  mM.

#### **CoaD as target for anti-infective development.**

Coenzyme A biosynthesis is essential for bacterial viability. CoaD,  
15 phosphopantetheine adenylyltransferase, catalyzes the fourth step in the pathway and was shown to be essential in both *E. coli* and *S. aureus*. There is no measurable homology between CoaD and the human PPAT enzyme, so the liability of poorly selective compounds is quite low. As CoaD is essential and conserved in gram-negative and gram-positive pathogens, inhibitors developed against this target will  
20 have broad-spectrum utility.

#### **Assays for measuring CoaD function**

CoaD will be expressed and purified using standard methodologies for bacterial expression and affinity tag-based purification. Two assay formats can be  
25 used to monitor enzymatic activity: the forward reaction and the reverse reaction.

The forward reaction assay was initially described for measuring the activity of the human PPAT activity in the PPAT/DPCK enzyme. The enzyme assay is carried out in 50 mM Tris (pH 8.0), 2 mM MgCl<sub>2</sub>, 5 mM ATP, 5-500  $\mu$ M 4'-phosphopantotheine, 7.5 mM NADH and enzyme (initially 0.1 – 1.0  $\mu$ g/ml). The  
30 production of PP<sub>i</sub> is detected using the protocol of O'Brien in which PP<sub>i</sub> production is coupled to the oxidation of NADH to NAD. This system requires the addition of 4 enzymes (PP<sub>i</sub>-dependent phosphofructokinase, aldolase, triosephosphate isomerase and glycerol-3-P dehydrogenase) to the basic reaction mix and presents the added issue of deconvolution, which limits the use of the assay as a primary screen.

The reverse direction assay is carried out also as a coupled assay to tie ATP production to NADP reduction following the method described by Lamprecht & Trautschold. The assay is set up in reaction buffer containing the following: 50 mM Tris (pH 8.0), 1 mM DTT, 2 mM MgCl<sub>2</sub>, 1 mM NADP, 5 mM glucose, 2 mM PP<sub>i</sub>, 0.1 mM dPCoA. Hexokinase (4 units) and glucose-6-phosphate dehydrogenase (1 unit) will be added to the assay as the coupling enzymes in addition to CoaD (initially 0.1 – 1 µg/ml). The assay is monitored at 340 nm. Deconvolution of hits is required with this assay, however with only 2 additional enzymes the task will be less cumbersome when compared to the forward assay described above.

#### Example 9

##### IDENTIFICATION OF THE GENE ENCODING DEPHOSPHOCO A KINASE (DPCK, YAC E, COAE) IN ALLOIOCOCCUS OTITIDIS

DephosphoCoA kinase (DPCK, YacE, CoaE) encoded by the *coaE* gene catalyzes the final step in Coenzyme A (CoA) biosynthesis. The final step in CoA biosynthesis is the phosphorylation of the 3'-hydroxyl group of dephospho-CoA to form CoA by dephosphocoenzyme A kinase (DPCK, YacE, CoaE).

The determination that the previously identified *yacE* gene encoded the dephosphocoenzyme A kinase activity was reported by Mishra *et al.* These authors previously determined that separate enzymes encode the phosphopantetheine adenylyltransferase (PPAT) and dephosphocoenzyme A kinase (DPCK) activity in *Corynebacterium ammoniagenes* in contrast to the eukaryotic enzymes in which the PPAT and DPCK activities are coupled. The *E. coli* gene, encoding a 25 kDa protein, was cloned based on the sequence of the *C. ammoniagenes* gene and found to be identical to the previously described *yacE* gene. The gene was designated *coaE* to follow existing nomenclature in *E. coli*. *CoaE* (*YacE*) was shown to be essential in *E. coli* through genetic footprinting. *CoaE* is widely distributed in bacteria. Homologue of this gene identified in *Alloiococcus otitidis* as described in Example 5 (Seq. ID No 93). The protein encoded by the gene is set forth in Seq. ID No. 94.



### Assays for measuring CoaE function

CoaE carries out the phosphorylation of dephosphocoenzyme A at the 3' hydroxyl group, consuming ATP, to form CoA. Dephosphocoenzyme A kinase activity is measured in a coupled reaction in which NADH oxidation to NAD is tied to ADP production. In this assay, the standard pyruvate kinase/lactose dehydrogenase coupling system is used to generate NAD in a 1:1 molar equivalent to the ADP produced by the test enzyme. NADH oxidation to NAD is monitored at 340 nm in a standard spectrophotometer. The following kinetic values were determined for CoaE:  $K_m(\text{ATP}) = 0.74 \text{ mM}$ ;  $K_m(\text{dephospho-CoA}) = 0.14 \text{ mM}$  (7).

The formation of CoA is monitored using a coupled enzyme system in which acetyl-CoA is formed in proportion to the amount of CoA in the assay. Three enzymes (phosphate acetyl transferase, citrate synthase and malate dehydrogenase) are added to the reaction that results in the formation of NADH from NAD, which is monitored at 340 nm.

### CoaE as a target for anti-infective development

Coenzyme A biosynthesis is essential for bacterial viability. CoaE, dephosphocoenzyme A kinase, catalyzes the final step in CoA synthesis and is shown to be essential by genetic footprinting in *E. coli*. A degree of homology between CoaE and the human DPCK enzyme has been noted, such that selectivity assays is necessary to determine a high therapeutic index for CoaE inhibitory compounds. CoaE is conserved in gram-negative and gram-positive pathogens and should have broad-spectrum utility in the clinic.

CoaE is expressed and purified using standard methodologies for bacterial expression and affinity tag-based purification. DephosphocoA kinase activity is monitored using a coupled enzyme system to tie ADP production to oxidation of NADH to NAD. The decay of absorbance at 340 nm will be the assay readout. The assay will be setup in the following buffer: 50 mM Tris (pH 8.5), 20 mM KCl, 10 mM  $\text{MgCl}_2$ , 10 mM ATP, 0.3 mM NADH and 0.4 mM phosphoenolpyruvate. The coupling enzymes: pyruvate kinase (10 U) and lactate dehydrogenase (4 U) will be added along with dephosphocoenzyme A kinase (initially 0.1- 1.0 ug/ml). The assay will be started by the addition of 0.4 mM dephosphocoenzyme A. In this assay system, the release of ADP is tied to the oxidation of NADH to NAD, and is monitored at 340 nm.

This assay is transferable to a high-density microtiter plate format and suitable for HTS.

#### EXAMPLE 10

5     **IDENTIFICATION OF DNAB AND PCRA, GENES ENCODING HELICASES IN *ALLOIOCOCCUS OTITIDIS***

10     Helicases unwind double-stranded DNA in a reaction that couple nucleotide binding and hydrolysis to strand unwinding. Their activity is required for a number of biological processes such as separation of the chromosome during replication, recombination and repair. Homologue of these genes were identified in *Alloiococcus otitidis* as described in Example 5 (Seq. ID No 15 and 99). The protein encoded by the gene is set forth in Seq. ID No. 16 and 100.

15     Due to the essential roles modulated by these molecules they represent an important target for antibacterial therapy. Homologs of *dnaB* and *pcrA* genes encoding helicases were identified as described in Example 5. A primary assay, which detects helicase function *in vitro*, is used to identify inhibitors of each enzyme and is described below.

20     Genes encoding DnaB and PcrA is obtained using polymerase chain reaction amplification of the genomic region encoding them. The genes is subcloned into a standard expression vector either containing an amino acid tag for ease of purification or not. The enzyme is then over-expressed in *Escherichia coli* and purified using a standard tag system.

25     Most helicases require a region of single-stranded DNA flanking the duplex region that it unwinds. As a result, providing a single stranded region to either the 3' or 5' end of a duplex allows for determination of the polarity of helicase unwinding. These types of experiments have demonstrated that PcrA and DnaB are 3'-5' and 5'-3' helicases, respectively. None the less, a convenient filtration assay has previously  
30     been described that is formatted for high-through-put screening of inhibitors of either enzyme, regardless of polarity. Assays (90 ul) contained 15 pM single-stranded M13 DNA to which a radiolabeled oligonucleotide had been annealed as a substrate for unwinding. Reactions are carried out in 96-well GF/C unfilter hydrophobic plates (Polyfiltronics Inc.) in 70 ul helicase buffer [20 mM Hepes (pH 7.6), 4 mM MgCl<sub>2</sub> 4

mM ATP, 100 ug/ml BSA, 5% glycerol and 2 mM DTT] and 10 ul of DMSO or compound. Reactions are initiated by adding 10 ul of purified helicase protein and are incubated for 1 hr at room temperature. 100 ul of 2X capture buffer containing silica beads [25% methanol, 3 M NaI, 0.03% NP-40, and 10% GlassFog beads (BIO101)] were added. The mixture was incubated for 30 min at room temperature. Plates are then washed 5X on a Bio-Tek instruments, Auto Washer EL403) with wash buffer (50% ethanol, 0.2% NP-40 and 50 mM NaCl). Scintillation fluid was added and plates are counted (Packard Topcount).

#### EXAMPLE 11

##### IDENTIFICATION OF DnaE, THE GENE ENCODING DnaE-POLYMERASE IN *ALLOIOCOCCUS OTITIDIS*

DnaE is an enzyme that catalyzes the DNA template directed polymerization of deoxyribonucleotides into deoxyribonucleic acid. The enzyme has been reported to modulate lagging strand synthesis at gram-positive replication forks. Functions for DnaE have been defined biochemically, in *Bacillus subtilis* and *Streptococcus pyogenes*. Homologue of this gene identified in *Alloiococcus otitidis* as described in Example 5 (Seq. ID No 75). The protein encoded by the gene is set forth in Seq. ID No. 76.

Because DnaE is an essential protein in gram-positive bacteria and has high homology to the gram-negative *dnaE*, which is an essential polymerase subunit of the DNA polymerase III holoenzyme, it serves as a good target for antibacterial drug discovery. A primary assay, which detects processive DnaE mediated DNA synthesis *in vitro*, is useful identify inhibitors of the enzyme and is described below.

The gene encoding DnaE I in *Alloiococcus otitidis* was identified as described in Example 5. Purification of DnaE DNA polymerase from *Alloiococcus*. The gene encoding DnaE is obtained using polymerase chain reaction amplification of the *dnaE* gene. The gene is subcloned into a standard expression vector either containing an amino acid tag for ease of purification or not. The enzyme is then over-expressed in *Escherichia coli* and purified using a standard tag system.

Because DnaE catalyzes the incorporation of single deoxyribonucleotides into DNA, the incorporation of radiolabelled deoxyribonucleotides into larger deoxyribonucleic acid molecules is monitored to measure activity of the enzyme. A

filtration assay has been previously described for *Streptococcus pyogenes* DnaE that uses filterplates containing DE81 filters to capture polymerized DNA. This assay is amenable to high-through-put screening format for DnaE. Assays contained 70 ng of 30-mer primed M13mp18 single stranded DNA as a template for replication. The  
5 reaction contained 3.3-300 ng of DnaE in 23.5  $\mu$ l of replication buffer [20 mM Tris-HCL (pH 7.5), 4% glycerol, 0.1 mM EDTA, 5 mM DTT, 2 mM ATP, 8 mM MgCl<sub>2</sub>, 40  $\mu$ g/ml BSA] and 60  $\mu$ M of both dGTP and dCTP. NaCl was added to the reaction mixture to a final concentration of 40 mM. DNA synthesis was initiated by the addition of 1.5  $\mu$ l of 1.5 mM dATP and 0.5 mM [ $\mu$ -<sup>32</sup>P]dTTP. Reactions were  
10 incubated at 37°C for various lengths of time and were quenched by adding an equal volume of 1% SDS and 40 mM EDTA. One-half of the terminated reaction was applied to DE81 filter paper and washed 3X with wash solution (0.3 M Ammonium formate and 0.01 M Sodium pyrophosphate). Filters were then placed in scintillation vials and 1 ml scintillation counting liquid was added. Radioactivity was counted  
15 using a scintillation counter.

## EXAMPLE 12

### IDENTIFICATION OF DNAG, THE GENE ENCODING PRIMASE IN *ALLOIOCOCCUS OTITIDIS*

20 DnaG is an enzyme that catalyzes the DNA template directed polymerization of ribonucleotides into ribonucleic acid *de novo*. Ribonucleic acid molecules that are synthesized by DnaG primase subsequently serve as primers for synthesis of the leading- and lagging-strands during chromosomal replication. Functions for DnaG have been defined biochemically, and the crystal structure of the RNA polymerase  
25 domain has been determined in *Escherichia coli*. Homologue of this gene identified in *Alloiococcus otitidis* as described in Example 5 (Seq. ID No 63). The protein encoded by the gene is set forth in Seq. ID No. 64.

Because DnaG primase plays an essential role in both leading- and lagging-strand synthesis during chromosomal replication, and DnaG has homologs in all  
30 prokaryotes but not eukaryotes, it serves as a good target for antibacterial drug discovery. A primary assay, which detects DnaG mediated RNA synthesis *in vitro*, can be used to identify inhibitors of the enzyme and is described below.

### Assay for the activity of DNA polymerase and identification of compounds that inhibit DnaG

The gene encoding DnaG is obtained using polymerase chain reaction amplification of the *dnaG* gene. The gene is subcloned into a standard expression vector either containing an amino acid tag for ease of purification or not. The enzyme is then over-expressed in *Escherichia coli* and purified using a standard tag system.

Because DnaG catalyzes the incorporation of single ribonucleotides into RNA, the incorporation of radiolabelled ribonucleotides into larger ribonucleic acid molecules is monitored to measure activity of the enzyme. A high-throughput scintillation proximity assay (SPA) assay, previously described for *E. coli* DnaG, is used to measure activity of DnaG activity in a coupled reaction with DnaB helicase. The assay, which was shown to work with DnaG alone, is used to screen for compounds that inhibit DnaG function. Assays are run in 96-well Packard Optiplate plates. First, 1 µl DMSO or test compound was added, followed by 20 µl of DnaG (208 nM) and 3.3 nM M13mp18 single-stranded DNA. Reactions are initiated by adding 10 µl of primase assay buffer [50 mM Tris-HCl (pH 7.5), 4% sucrose, 8 mM DTT, 5 mM MgCl<sub>2</sub>, 40 µg/ml BSA, 0.1 µg/ul Rifampicin, 25 U/ml RNA guard, 100 µM GTP, 100 µM UTP, 3 µM CTP, 1 mM ATP] and 0.4 µCi [<sup>3</sup>H]CTP. Reactions are incubated at 30°C for 30 min. Next, a suspension of 50 µl of 2.5 mg/ml PVT-PEI SPA beads (Amersham; prepared in 0.3 M NaCitrate, pH 3.0) were added. Plates were read after 1 hr on a Topcount instrument (Packard).

### EXAMPLE 13

#### DNAN, DNAX, HOLA, HOLB, AND POLC, THE GENES ENCODING THE SUBUNITS OF ALLOIOCOCCUS OTITIDIS DNA POLYMERASE III HOLOENZYME: BETA (β), TAU (τ), DELTA (Δ), DELTA' (Δ') AND POLC.

DNA polymerase III holoenzyme is an enzyme complex comprised of multiple highly conserved subunits that catalyzes the DNA template directed polymerization of deoxyribonucleotides into deoxyribonucleic acid. In gram positive organisms the holoenzyme is composed of a polymerase subunit, PolC, and accessory proteins. The accessory proteins act in a coordinated manner to clamp the polymerase tightly to the DNA template allowing the polymerase to synthesize DNA with high speed and

processivity. Homologue of these genes identified in *Alloiococcus otitidis* are described in Example 5 (Seq. ID Nos. 21, 105, 79, 103, and 105 respectively). The protein encoded by the gene is set forth in Seq. ID No. 22, 106, 80, 104 and 106 respectively).

5 Functions for the individual subunits have been defined biochemically and interactions between them have now been deduced structurally by crystallographic analysis of the enzyme from *Escherichia coli*. Tau interacts directly with both delta and delta' to form a clamp loader complex. Upon binding ATP the complex undergoes a conformational change altering an interaction between delta and delta',  
10 which allows delta to subsequently interact with the beta-clamp. The beta-clamp is a ring-shaped homomultimer assembly that can be opened by delta and placed onto a primed DNA template. ATP hydrolysis results in closing the clamp around DNA and dissociation of the clamp-loading complex. PolC then couples with the beta clamp to form a highly processive polymerase.

15 Because DNA polymerase III holoenzyme is comprised of multiple subunits, the opportunity exists to inhibit its activity at a number of different sites. A primary assay, which detects processive DNA synthesis *in vitro*, can be used to identify inhibitors of the enzyme and is described below. Deconvolution of inhibitors, based on either activity of physical interaction, follow the primary assay.

20

#### **Assay for the activity of DNA polymerase**

Purification of DNA polymerase III holoenzyme subunits from *Alloiococcus*. Genes encoding the subunits of DNA polymerase is obtained using polymerase chain reaction (PCR) amplification of the genomic region encoding them. The genes  
25 are subcloned into a standard expression vector either containing an amino acid tag for ease of purification or not. The enzyme is then over-expressed in *Escherichia coli* and purified using a standard tag system.

Because DNA polymerase III catalyzes the incorporation of single deoxyribonucleotides into DNA, the incorporation of radiolabeled deoxynucleotides  
30 into larger deoxyribonucleic acid molecules is monitored to measure activity of the enzyme. A filtration assay is previously described for *Streptococcus pyogenes* DNA polymerase III that uses filterplates containing DE81 filters to capture polymerized DNA (2). This assay is amenable to high-through-put screening format. Assays

contained 70 ng of 30-mer primed M13mp18 single stranded DNA as a template for replication. The reaction contained 43 ng of  $\beta$  and 140 ng of PolC- $\tau$  complex in 23.5  $\mu$ l of replication buffer (20 mM Tris-HCL (pH 7.5), 4% glycerol, 0.1 mM EDTA, 5 mM DTT, 2 mM ATP, 8 mM  $MgCl_2$ , 40  $\mu$ g/ml BSA, and 60  $\mu$ M of both dGTP and dCTP. DNA synthesis was initiated by the addition of 1.5  $\mu$ l of dATP and [ $\mu$ - $^{32}P$ ]dTTP. Reactions were incubated at 37°C for various lengths of time and were quenched by adding an equal volume of 1% SDS and 40 mM EDTA. One-half of the terminated reaction was applied to DE81 filter paper and washed 3X with wash solution (0.3 M Ammonium formate and 0.01 M Sodium pyrophosphate). Filters were then placed in scintillation vials and 1 ml scintillation counting liquid was added. Radioactivity was counted using a scintillation counter.

Compounds inhibiting PolC subunit is identified by modifying the above reaction to include only the PolC subunit and using 2.5  $\mu$ g activated calf thymus DNA as a substrate, instead of singly-primed M13mp18 DNA, as previously described. Several techniques are utilized to determine the interaction of inhibitors with individual subunits. These have been described in the literature and include the following: (1) Nuclear magnetic resonance and capillary electrophoresis.

20

#### EXAMPLE 14

##### ERA GTPASE IN ALLOIOCOCCUS OTTIDIS

The *era* (*E. coli* Ras) gene was initially identified while sequencing around the *rnc* gene; *era* lies downstream of *rnc*. While a function for *era* has yet to be determined, conditional (temperature sensitive) mutants revealed that the product of the *era* gene, Era, is essential for *E. coli* viability. A hint as to an *in vivo* function for Era was uncovered when a suppressor of a *dnaG* (primase) allele was found to map in the *era* coding sequence and a second suppressor, which mapped upstream of the *era* open reading frame, affected expression of *era*. These data suggest that Era could play one or more roles in DNA replication, regulation of primase activity or otherwise effect cell cycle progression. More recent data has confirmed that the *era*1 mutant causes a defect in cell growth at the two-cell stage and delays cell division. Moreover, Britton et al demonstrated that cell division was coupled with the level of

30

Era in the cell: division arrest, through reduction in Era levels, is reversed when Era levels return to threshold amount. A current model suggests that Era acts as a checkpoint regulator in the bacterial cell cycle. Era is a GTP-binding protein with GTPase activity, a threshold level of functional/activated Era may be required to initiate septation.

Era is associated with additional cellular functions, specifically translation, as Era specifically interacts with the translation machinery. *E. coli* Era binds both 16S rRNA and the 30S ribosomal subunit; whereas, the *S. pneumoniae* 16S rRNA co-purifies with Era. A putative RNA binding "KH motif" has been identified in the carboxyl-terminal domain. The RNA binding activity is critical to Era cellular function as mutation of the putative RNA binding region of the *S. pneumoniae* Era prevents complementation of an *E. coli era* mutant strain. Homologue of this gene identified in *Alloicoccus otitidis* as described in Example 5 (Seq. ID No 65). The protein encoded by the gene is set forth in Seq. ID No. 66.

#### Nucleotide binding

Filter-binding assays are utilized to demonstrate nucleotide-binding specific to GTP and not UTP, CTP or ATP. Both GTP and GDP (unlabeled) were capable of inhibiting  $\alpha^{32}\text{P}$ -GTP binding. The Kd for GTP and GDP binding were reported to be 5.5 and 1.0  $\mu\text{M}$ , respectively.

A large number of GTP-binding proteins have been studied and all members of the family contain three regions of highly homologous amino acid residues that define a GTP-binding pocket. Era contains well-conserved regions defining the so-called G1 (G/AXXXGKT/S: residues 15-22), G3 (DXXG: residues 62-65) and G4 (NKXD: residues 124-128) consensus sequences. The G2 domain (residues 33-38, see below), located between G1 and G3, is generally more variable.

#### GTPase activity

Purified Era showed a significant GTPase activity, which is inhibitable by GTP or GDP but not by UTP, CTP, ATP or ADP. The maximum hydrolysis rate is measured at 9.8 mmol GTP hydrolyzed/min/mol Era. The Km was found to be 9  $\mu\text{M}$ .

It should be noted that Sullivan *et al* demonstrated, using mant (*N*-methyl-3'-O-anthraniloyl) labeled GTP and GDP, very rapid exchange kinetics for guanine



nucleotide binding. Era exchanges guanine nucleotides 10-fold more rapidly than the GTP hydrolysis rate suggesting that guanine nucleotide binding and release should be considered as a regulatory point in addition to the more well-studied hydrolysis step.

5

### Autophosphorylation

When  $\gamma^{32}\text{P}$ -GTP is used as a substrate for the GTPase activity, Era is phosphorylated. The autophosphorylation reaction is specific for GTP, as incubation with  $\gamma^{32}\text{P}$ -ATP did not result in phosphorylation of Era. Moreover,  $\alpha^{32}\text{P}$ -GTP is not a  
10 suitable substrate for detection of Era autophosphorylation. Tryptic digestion and HPLC were utilized to resolve the sites(s) of phosphorylation. Using  $\gamma^{32}\text{P}$ -GTP as a substrate the major radioactive peak contained the tryptic peptide, ISITSR, corresponding to Era residues 33-38 and containing 3 potential phosphorylation sites. Mutagenesis of both Thr-36 and Ser-37 to alanine abolished enzymatic  
15 activity. However, individual alanine substitutions at either site had no effect on Era function. The autophosphorylation site is located in the so-called G2 domain of Era.

### Suitability of target for anti-infective development

Era is an essential protein for bacterial viability. Knock-down mutations as well  
20 as conditional-lethal alleles revealed that Era function is required for cytokinesis. An additional phenotype of the Era-depleted strains is an aberrant response to temperature induced stress. This target is novel and may well lead to the identification of new classes of anti-infectives. The widespread distribution of Era homologues in both gram-negative and gram-positive pathogens suggests that  
25 broad-spectrum agents could result from an effort to define Era inhibitory compounds.

### Assays for measuring Era function

#### 30 NUCLEOTIDE BINDING ASSAYS

Era binding to nucleotide is monitored by a simple filter-binding assay. Era (1-5  $\mu\text{g}$ ) is incubated with  $\alpha^{32}\text{P}$ -GTP (0.2  $\mu\text{Ci}$ ) in a buffer consisting of 100 mM Tris (pH 7.5), 10 mM  $\text{MgCl}_2$ , 0.2% NP-40, 0.2 mg/ml BSA for 30 minutes at 32°C. A

portion of the reaction mix is spotted on nitrocellulose membrane, washed (50 mM Tris (pH 7.5), 5 mM MgCl<sub>2</sub>, 1 mM DTT) and dried. The membrane is then exposed to X-ray film. Alternatively, the spots are excised and counted. This assay is directly amenable to HTS using filter plates.

5

#### GTPASE ACTIVITY ASSAY

The GTP hydrolytic activity of Era is monitored using thin-layer chromatography. Era and  $\alpha^{32}\text{P}$ -GTP is incubated in 50 mM Tris (pH 7.5), 5 mM MgCl<sub>2</sub>, 0.1% NP-40, 0.2 mg/ml BSA for 30 minutes at 37°C. An aliquot of the  
10 reaction is placed on PEI cellulose and the strip developed with 0.5 M KH<sub>2</sub>PO<sub>4</sub>, 1.0 M NaCl (pH 3.7). The spots conforming to GDP and GTP are identified by UV shadowing, excised and counted. This assay represents an acceptable secondary/confirmatory assay.

Alternatively, the hydrolysis of  $\gamma^{32}\text{P}$ -GTP is monitored by assaying for  
15 liberated P<sub>i</sub>. Obg and  $\alpha^{32}\text{P}$ -GTP is incubated in 50 mM Tris (pH 8.5), 1.5 mM MgCl<sub>2</sub>, 0.1 mM EDTA, 100 mM KCl, 10% glycerol for 30 minutes to 3 hours at 37°C. The reaction will be stopped by the addition of a slurry of charcoal in 1 mM Kpi (pH 7.5), which selectively binds the GTP and GDP. The liberated P<sub>i</sub> in the supernatant is monitored by Cerenkov counting. Free P<sub>i</sub> can also be monitored with the Malachite  
20 Green reagent.

#### AUTOPHOSPHORYLATION ASSAY

Era autophosphorylation is monitored by incubating Era with  $\gamma^{32}\text{P}$ -GTP in 50 mM morpholinopropane sulphate (pH 6.8), 5 mM MgCl<sub>2</sub>, 1 mM DTT at 37°C (14).  
25 Samples are analyzed following separation on SDS polyacrylamide gels, drying the gel and exposure to film. This assay represents an acceptable secondary/confirmatory assay for Era activity.

## EXAMPLE 15

FmHB(FEMX) GENES IN ALLOIOCOCCUS OTITIDIS

5       The *femA*, *femB*, and *fmbB(femX)* genes have been shown to be essential for incorporation of glycine into the side chain of peptidoglycan precursors in *Staphylococcus aureus*. The *femAB* locus was initially identified as a factor essential for methicillin resistance (*fem*) based on random insertional inactivation of chromosomal genes and a screen for reduced expression of resistance mediated by  
10       the penicillin binding protein 2A (PBP2A). Inactivation of *femA* or *femB* was subsequently reported to prevent incorporation of glycine residues at positions 2 to 5 or positions 4 to 5 of the penta-glycine cross bridge since mucopeptides cross-linked by one or three glycine residues were detected in the corresponding mutants. Inactivation of *fmbB*, formerly *femX*, is lethal, but the construction of a mutant  
15       conditionally expressing *fmbB* under the control of a xylose-inducible promoter showed that the gene was essential for synthesis of branched peptidoglycan precursors. These studies show that the *fem* gene products were required for incorporation of glycine at positions 1 (FmbB), 2 and 3 (FemA), and 4 and 5 (FemB) of the cross bridge, although the catalytic activity of the proteins has not been directly  
20       assessed. Similarly, inactivation of two *fmbB* homologues in *Streptococcus pneumoniae*, designated *murM (fibA)* and *murN (fibB)*, reduced addition of L-Ala or L-Ser to the -amino group of L-Lys and subsequent addition of a second L-Ala residue, respectively. Overall, disruption of the *murMN* operon reduced the proportion of  
25       branched peptide stems in the peptidoglycan from 89 to 33%. In contrast to what occurs in *S. aureus*, direct cross-linking of L-Lys to D-Ala occurs in *S. pneumoniae*, and the *murMN* operon was accordingly reported to be unessential.

      BLAST analysis of *Alloiococcus otitis* genome revealed an ORF similar to *femX* of *Weissella viridescens*, and *fmbB* of *S. aureus*. It suggests that in *Alloiococcus otitis* there is an enzyme with similar to FmbB function. Homologue of  
30       this gene identified in *Alloiococcus otitis* is described in Example 5 /Table 4 (Seq. ID No 97). The protein encoded by the gene is set forth in Seq. ID No. 98.

### Assays for measuring FmhB function

There are no *in vitro* biochemical assays to test enzymatic activity of *S. aureus* FmhB because the reaction occurs at the membrane-bound lipid II precursor GlcNAc-( $\beta$ -1,4)-*N*-acetylmuramic acid(-L-Ala-D-iGln-L-Lys-D-Ala-D-Ala)-pyrophosphoryl-undecaprenol.

Lipid II is a minor component of bacterial cell membrane which is detected by thin-layer chromatography separation of presolubilized membranes supplied with the cytoplasmic precursors, UDP-*N*-acetylmuramyl-pentapeptide (UDP-MurNac-pentapeptide) and [ $^{14}$ C]UDP-*N*-acetylglucosamine ([ $^{14}$ C]UDP-GlcNAc).

The *in vitro* biosynthesis of branched lipid II of *S. aureus* requires whole-cell membranes, cytoplasmic PG precursors, glycine ( $^{14}$ C labeled for detection of reaction products), purified tRNA, and an intracellular fraction that contains tRNA-activating enzymes. Therefore, the *in vitro* assay of *S. aureus* FmhB is a tedious procedure.

One way to facilitate this procedure is to use *Weissella viridescens* FemX or *E. faecalis* UDP-MurNac-pentapeptide:L-alanine ligase. Recombinant *Weissella viridescens* FemX and *E. faecalis* UDP-MurNac-pentapeptide:L-alanine ligase were purified, and their *in vitro* activity was demonstrated. The distinctive feature of these enzymes is that they catalyze the addition of a branching amino acid (Ala) to the cytoplasmic cell wall precursor UDP-MurNac-pentapeptide.

Other bacteria for which the biosynthesis of Gly-containing branched UDP-MurNac-hexapeptide in cytoplasm was shown are *Streptomyces lividans* and *Streptomyces hydroscopicus*, although the enzymes were not isolated and their ligase activity remain to be demonstrated.

These new data open an opportunity to develop an assay to detect the activity of FmhB(FemX) by using cytoplasmic UDP-MurNac-pentapeptide. Products of the reaction are detected by HPLC. HPLC separation of precursors are performed by the method of Flouret et al. The precursors are separated by reverse-phase HPLC on a  $\mu$ Bondapak C<sub>18</sub> column (3.9 by 300 mm; Waters) in 50 mM ammonium formate (pH 3.9) at a flow rate of 0.5 ml/min. The elution of precursors is monitored at a wavelength of 254 nm.

## EXAMPLE 16

FOLA- DIHYDROFOLATE REDUCTASE (DHFR)

5       The *Alloicoccus* ORF-1863 encodes a homolog of *S. aureus* dihydrofolate reductase that catalyzes the NADPH-dependent conversion of dihydrofolate to tetrahydrofolate, one of the steps in bacterial folate biosynthesis. Homologue of this gene identified in *Alloicoccus otitidis* is described in Example 5/Table 4 (Seq. ID No 55). The protein encoded by the gene is set forth in Seq. ID No. 56.

10

**FOLA as a target for anti-infective development**

Folate is an essential cofactor in many important metabolic processes in bacteria, such as purine, pyrimidine, amino acid and pantothenate biosynthesis. Unlike mammalian cells, bacteria are unable to utilize exogenous folate derivatives, and therefore must synthesize folate *de novo*. Bacterial folate biosynthesis occurs via two converging pathways, the non-essential para-amino-benzoate (PABA) synthesis pathway, and synthesis of the pterin precursor, to which pABA is subsequently attached to form the folate precursor. Bacterial DHFRs are essential for viability and well conserved across all bacterial species. Although bacterial DHFR shares  
15       similarity with human DHFR, selective inhibitors against bacterial DHFR have been identified in the past such as trimethoprim which specifically blocks the bacterial DHFR step. Thus DHFR still remains an attractive target for development of broad-spectrum antibacterial agents.

25       **Assays for measuring DHFR activity**

DHFR activity is monitored spectrophotometrically, recording the change of absorbance at 340 nm due to the equimolar consumption of NADPH in the course of dihydrofolate substrate reduction. DHFR (10 ng) is preincubated in reaction buffer containing 50 mM 2-(N-morpholino)ethanesulfonic acid, 25 mM Tris-HCl, 25 mM  
30       ethanolamine, and 100 mM NaCl at pH 7.5 for 3 minutes. The reaction is started by addition of 0.5-10  $\mu$ M 7,8-dihydrofolate. The amount of processed substrate is calculated from the decrease of absorbance at 340 nm due to oxidation of NADPH ( $\epsilon=11800 \text{ M}^{-1}\text{cm}^{-1}$ ) to  $\text{NADP}^+$ .

## EXAMPLE 17

**FOLB- DIHYDRONEOPTERIN ALDOLASE (DHNA)**

5           The *Alloicoccus otitidis* ORF-959 encodes a homolog of *S. aureus*  
dihydroneopterin aldolase that catalyzes the conversion of 7,8-dihydroneopterin to 6-  
hydroxymethyl-7,8-dihydropterin, one of the early steps in bacterial folate  
biosynthesis. Homologue of this gene identified in *Alloicoccus otitidis* is described in  
Example 5/Table 4 (Seq. ID No 31). The protein encoded by the gene is set forth in  
10   Seq. ID No. 32.

**FOLB as a target for anti-infective development**

Folate is an essential cofactor in many important metabolic processes in  
bacteria, such as purine, pyrimidine, amino acid and pantothenate biosynthesis.  
15   Unlike mammalian cells, bacteria are unable to utilize exogenous folate derivatives,  
and therefore must synthesize folate *de novo*. Bacterial folate biosynthesis occurs via  
two converging pathways, the non-essential para-amino-benzoate (*pABA*) synthesis  
pathway, and synthesis of the pterin precursor, to which *pABA* is subsequently  
attached to form the folate precursor. Enzymes that catalyze steps in the folate  
20   biosynthesis pathway are essential and well conserved across all bacterial species,  
and those that act in early steps such as FolB have no direct homologs in mammals.  
Thus FolB becomes an attractive target for development of broad-spectrum  
antibacterial agents.

**25   Assays for measuring FOLB activity**

FolB (DHNA) 7,8-dihydroneopterin aldolase activity is monitored individually  
or in conjunction with downstream enzymes in folic acid biosynthesis pathway (FolK  
and Sul).

FolB activity is monitored directly by HPLC assay. FolB substrate (7,8-  
30   dihydro-D-neopterin) is commercially available from Schircks Laboratories  
(Switzerland). FolB (0.5 µg) is preincubated in reaction buffer containing 50 mM Tris-  
HCl (pH 8.0), 50 mM KCl, 0.1 mg/ml BSA, 2.5 mM dithiothreitol for 5 min. Reaction  
is started by addition of stock solution of 7,8-dihydro-D-neopterin in DMSO (100 µM

final concentration). Reaction is terminated by addition of 1/3 of reaction volume of 1% I<sub>2</sub>, 2% KI in 1M HCl with subsequent incubation at room temperature for 5 minutes. Quenched reaction will be applied directly to HPLC. Oxidized starting material and reaction products are efficiently separated on ODS (C18) column.

5 Reaction components are detected and quantified by analysis of UV absorbance at 254 nm, or fluorescence (excitation at 365 nm; emission at 446 nm).

FolB activity are also monitored in the coupled assay with FolK (HPPK) and Sul (DHPS) enzymes. FolB activity is measured by detection of radioactive dihydropteroate formation as described in FolK and Sul assays, under conditions of  
10 excess of the later enzymes. FolB enzyme and substrate 7,8-dihydro-D-neopterin are added to the described assay to replace the 6-hydroxymethyl-7,8-dihydropterin (FolK substrate).

#### EXAMPLE 18

##### 15 FOLC- DIHYDROFOLATE SYNTHASE (DHFS)

The *Alloiococcus otitidis* ORF-956 and ORF-528 both encode a homolog of *B. subtilis* dihydrofolate synthase that catalyzes the conversion of 7,8-dihydropteroate and glutamate to dihydrofolate, one of the steps in bacterial folate biosynthesis [.

20 Homologue of this gene identified in *Alloiococcus otitidis* as described in Example 5 (Seq. ID Nos. 29 and 23). The protein encoded by the gene is set forth in Seq. ID Nos. 30 and 24.

#### **Use of FolC as a target for anti-infective development**

25 Folate is an essential cofactor in many important metabolic processes in bacteria, such as purine, pyrimidine, amino acid and pantothenate biosynthesis. Unlike mammalian cells, bacteria are unable to utilize exogenous folate derivatives, and therefore must synthesize folate *de novo*. Bacterial folate biosynthesis occurs via two converging pathways, the non-essential para-amino-benzoate (pABA) synthesis  
30 pathway, and synthesis of the pterin precursor, to which pABA is subsequently attached to form the folate precursor. Enzymes that catalyze steps in the folate biosynthesis pathway are essential, and are well conserved across all bacterial species. Bacterial FolC appears to be a bifunctional enzyme possessing both

dihydrofolate synthase (DHFS) activity and folyl-polyglutamate synthetase (FPGS) activity, which are probably mediated through different sites of the protein. The bacterial DHFS activity but not the FPGS activity is essential for viability. Although bacterial FolC shares similarity with human FPGS, the human enzymes apparently lack DHFS activity and display a folate substrate specificity quite distinct from that of bacterial enzymes. Thus targeting bacterial FolC/DHFS activity selectively might lead to identification of broad-spectrum antibacterial agents.

#### Assays for measuring FolC activity

FolC (DHFS) 7,8-dihydrofolate synthase activity in the presence or absence of antimicrobial compounds or putative inhibitory compounds are monitored by several methods.

In one method, FolC activity is monitored directly by simple HPLC assay. FolC substrate (7,8-dihydropteroic acid) is commercially available from Schircks Laboratories (Switzerland). FolC (15 ng) is added to reaction mix, containing 10 mM glutamate, 5 mM ATP, 50 mM Tris-HCl (pH 8.0), 20 mM Mg<sub>2</sub>Cl, 50 mM KCl, 0.1 mg/ml BSA, 5 mM dithiothreitol. Reaction is started by addition of stock solution of 7,8-dihydropteroic acid in DMSO (10  $\mu$ M final concentration). Reaction is terminated by addition of equal volume of 8M Guanidinium hydrochloride. Stopped reaction is applied directly to HPLC. Starting material and reaction products are efficiently separated on ODS (C18) column. Reaction components are detected and quantified by analysis of UV absorbance at 254 nm, or fluorescence (excitation at 280 nm; emission at 420 nm).

In another method, the FolC activity monitoring is by detection of ADP accumulation. ADP is released in the amount equimolar to the amount of the product formed. ADP detection is performed by coupling its conversion to ATP by pyruvate kinase in the presence of phospho(enol)pyruvate producing pyruvate. Lactate dehydrogenase reduces pyruvate to S-lactate in the presence of NADH. Course of reaction is monitored by decrease in absorbance at 340 nm due to oxidation of NADH ( $\epsilon=6220 \text{ cm}^{-1}\text{M}^{-1}$ ) to NAD<sup>+</sup>. Reaction conditions are as following: 5 mM dithiothreitol, 5 mM ATP, 380  $\mu$ M NADH, 10 mM glutamate, 2 mM phospho(enol)pyruvate, 50 mM KCl, 20 mM Mg<sub>2</sub>Cl, 50 mM Tris-HCl, 50  $\mu$ g of



pyruvate kinase, 50 µg of S-lactate dehydrogenase. Reaction is started by addition 7,8-dihydropteroic acid in DMSO (10 µM final concentration).

In yet another method, FolC activity is monitored through detection of inorganic phosphate release. Amount of inorganic phosphate in solution is quantified by:

- (i) its conversion by purinenucleoside phosphorylase leading to phosphorylation of MESG. Later assay kit is available from Molecular Probes as EnzCheck™ Phosphate Assay Kit;
- (ii) its reaction with Malachite Green reagent; and
- (iii) detecting the release of radioactive inorganic phosphate in reaction with  $\gamma$ -<sup>33</sup>P-labeled ATP following the absorption of unprocessed ATP by charcoal.

First method is applied in rate-based assay format; the later two in end-point format. Reaction conditions are similar to the ones described in HPLC-based assay.

#### EXAMPLE 19

##### FOLK- 6-HYDROXYMETHYL-7, 8-DIHYDROPTERIN PYROPHOSPHOKINASE (HPPK)

The *Alloiooccus otitidis* OFR-961 (Seq. ID No. 33) encodes a homolog of *S. aureus* 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase that catalyzes pyrophosphoryl transfer from ATP to 6-hydroxymethyl-7,8-dihydropterin, one of the early steps in bacterial folate biosynthesis. The protein encoded by this ORF is set forth in Seq. ID No. 34. (see Example 5/Table 4).

##### **Use of FolK as a target for anti-infective development**

Folate is an essential cofactor in many important metabolic processes in bacteria, such as purine, pyrimidine, amino acid and pantothenate biosynthesis. Unlike mammalian cells, bacteria are unable to utilize exogenous folate derivatives, and therefore must synthesize folate *de novo*. Bacterial folate biosynthesis occurs via two converging pathways, the non-essential para-amino-benzoate (*pABA*) synthesis pathway, and synthesis of the pterin precursor, to which *pABA* is subsequently attached to form the folate precursor. Enzymes that catalyze steps in the folate biosynthesis pathway are essential and well conserved across all bacterial species, and those that act in early steps such as FolK have no direct homologs in mammals.

Thus FolK is an attractive target for the development of broad-spectrum antibacterial agents.

#### Assays for measuring FolK activity

5 FolK (HPPK) 7,8-dihydroxymethylpterin-pyrophosphokinase activity is monitored individually or in conjunction with downstream enzyme in folic acid biosynthesis pathway.

10 FolK activity is monitored directly by HPLC assay. FolK substrate (7,8-dihydro-6-hydroxymethylpterin) is commercially available from Schircks Laboratories (Switzerland). FolK is preincubated in reaction buffer containing 50 mM Tris-HCl (pH 8.0), 50 mM KCl, 20 mM MgCl<sub>2</sub>, 5 mM ATP, 0.1 mg/ml BSA, 2.5 mM dithiothreitol. Reaction is started by addition of stock solution of 7,8-dihydro-6-hydroxymethylpterin in DMSO (100 μM final concentration). Reaction is terminated by addition of equal  
15 volume of 8M Guanidinium hydrochloride and applied directly on HPLC. Starting material and reaction products are efficiently separated on ODS (C18) column. Reaction components are detected and quantified by analysis of UV absorbance at 254 nm.

20 FolK activity is monitored by end-point assay coupled with excess of Sul enzyme. Activity is calculated from quantification of the radioactivity incorporated in final product (7,8-dihydropteroate).

#### EXAMPLE 20

##### ALLOIOCOCCUS OTITIDIS ENCODED FOLP (SUL)- DIHYDROPTEROATE SYNTHASE (DHPS)

25 The *Alloioococcus otitidis* ORF-1811 (Seq. ID No. 53) encodes a homolog of *B. subtilis* dihydropteroate synthase that catalyzes the condensation of pABA (para-aminobenzoic acid) with 6-hydroxymethyl-7,8-dihydropterin pyrophosphate, one of the early steps in bacterial folate biosynthesis. The polypeptide encoded by this ORF  
30 is set forth in Seq. ID No. 54. (see Example 5/Table 4)

#### FOLP AS A TARGET FOR ANTI-INFECTIVE DEVELOPMENT

Folate is an essential cofactor in many important metabolic processes in bacteria, such as purine, pyrimidine, amino acid and pantothenate biosynthesis. Unlike

mammalian cells, bacteria are unable to utilize exogenous folate derivatives, and therefore must synthesize folate *de novo*. Bacterial folate biosynthesis occurs via two converging pathways, the non-essential para-amino-benzoate (*p*ABA) synthesis pathway, and synthesis of the pterin precursor, to which *p*ABA is subsequently  
5 attached to form the folate precursor. Enzymes that catalyze steps in the folate biosynthesis pathway are essential and well conserved across all bacterial species, and those that act in early steps such as FolP (Sul) have no direct homologs in mammals. In fact, dihydropteroate synthase (FolP or Sul) is the target for known antibiotics sulfonamides which are competitive inhibitors of FolP/Sul as *p*ABA  
10 analogues. Thus FolP (Sul) still remains an attractive target for development of broad-spectrum antibacterial agents.

#### **Suitable assays for measuring FolP/Sul activity**

Sul (DHPS) 6-hydroxymethyl-7,8-dihydroneopteroate synthase activity is  
15 monitored individually or in conjunction with upstream enzymes in folic acid biosynthesis pathway (FolB and/or FolK).

DHPS activity is monitored directly by counting the amount of radioactivity incorporated in 6-hydroxymethyl-7,8-dihydroneopteroate when using radioactively  
20 labeled *p*-aminobenzoic acid (*p*ABA). Final product is separated from unreacted *p*ABA by thinlayer chromatography, paper chromatography or on HPLC equipped with radioactivity detector. DHPS substrate (6-hydroxymethyl-7,8-dihydropterin pyrophosphate) is not commercially available, but is quantitatively synthesized in one step from its oxidized precursor available from Schircks Laboratories (Switzerland). DHPS (20 ng) is added in reaction buffer containing 50 mM Tris-HCl, pH 8.0, 20 mM  
25 MgCl<sub>2</sub>, 0.1 mg/ml BSA, 5 mM dithiothreitol and 0.5 – 10  $\mu$ M PABA. Reaction is started by addition of stock solution of substrate (6-hydroxymethyl-7, 8-dihydropterin pyrophosphate, 0.05 - 1  $\mu$ M final concentration). Reaction is terminated by acidification of reaction volume with addition of equal volume of citrate/phosphate or ammonium acetate/acetate buffer, pH 4 containing excess of unlabelled *p*ABA.  
30 Quenched reaction is separated by chromatography and the amount of formed product calculated.

DHPS activity is determined in coupled assay with excess of FolB and FolK enzymes. The advantage of coupled assay is that it makes it possible to use

### EXAMPLE 21

The *Alloiococcus otitidis* ORF-2489 (Seq. ID No. 85) encodes a homolog of *E. faecalis* FtsA, one of the essential components of bacterial cell division. The “fts” stands for filamentation temperature sensitive and has been assigned to most bacterial cell division genes due to the fact that these genes were generally discovered by the isolation of conditional mutants that form filaments at nonpermissive temperature . The *ftsA* allele was first isolated and identified in *E. coli* by Ricard and Hirota in 1973, and mapped along with *ftsZ* in 1980. The protein encoded by this ORF is set forth in Seq. ID No. 86. (see Example 5/Table 4)

Bacterial cell division requires formation of a septum at mid-cell that begins with the polymerization of FtsZ into a ring structure at the nascent division site. FtsZ, another key component of bacterial septation is the first known protein to localize to the division site. In *E. coli*, shortly after the formation of the FtsZ ring, FtsA and ZipA (another key division component present only in gram-negative bacteria) [7] are independently recruited to the septal ring, most likely through their direct interaction with FtsZ. Subsequent assembly of other division components at the septum requires FtsA as well as FtsZ.

Like FtsZ, FtsA homologs are present and highly conserved in almost all eubacteria. FtsA is essential for cell division and its deletion leads to impaired cell division and sporulation defect. In addition, *E. coli* cells have to maintain critical ratio of FtsA to FtsZ in order for proper cell division to occur. FtsA belongs to the actin/DnaK/sugar kinase family of proteins. In *B. subtilis*, FtsA acting as a dimer not only binds ATP but also hydrolyzes ATP. As briefly stated above, *in vivo* and *in vitro* evidence have demonstrated that FtsA and FtsZ from various bacterial species

directly interact. Taken all together, targeting at FtsA especially at its interaction with FtsZ might lead to identification of broad-spectrum antibacterial agents.

#### Assays for measuring FtsA activity

5

ATPase activity of FtsA is assayed by following the formation of  $^{32}\text{P}$ i from  $[\gamma\text{-}^{32}\text{P}]\text{-ATP}$ . The reaction mixture containing 50 mM Tris-HCl (pH7.2), 50 mM potassium acetate, 1 mM DTT, 10 mM  $\text{MgCl}_2$  and different concentrations of  $[\gamma\text{-}^{32}\text{P}]\text{-ATP}$  is incubated for 5 minutes at 37°C. The reaction is started by addition of 50 nM purified FtsA of *Alloicoccus*. The reaction is stopped with 1.5% ammonium molybdate in 0.5N sulfuric acid, and the radioactive Pi extracted into isoamyl alcohol and counted.

10

Interaction between FtsA and FtsZ is detected quantitatively using yeast two-hybrid system as described. Briefly, *Alloicoccus ftsZ* is cloned into yeast two-hybrid bait vector pLexA (Clontech) to generate a LexA-FtsZ fusion with DNA-binding property. *Alloicoccus ftsA* is cloned into the target vector pB42AD (Clontech) to fuse FtsA to the activating domain. Both plasmids are then transformed into a *Saccharomyces cerevisiae* strain containing a *lacZ* reporter under the control of multiple LexA operators.  $\beta$ -Galactosidase activity is determined to quantify relative strength of FtsA-FtsZ interaction.

20

#### EXAMPLE 21

##### ALLOIOCOCCUS OTITIDIS ENCODED FILAMENTATION TEMPERATURE SENSITIVE GENE Z (FtsZ)

25

FtsZ is an essential protein that forms a cytokinetic ring (Z-ring) that drives cell division in bacteria. FtsZ has been identified in most prokaryotic species with the exception of *Chlamidia*, a *Ureaplasma* species and *Crenarchaea*. FtsZ and Z-ring formation are most extensively studied in *E. coli*. FtsZ is an abundant cytoplasmic protein which is present at  $\sim 10^4$  copies per cell, and is the first protein to be localized to the division site. Z-ring is required throughout septation and directs the ingrowth of septum in part by recruiting other cell division protein to the division site. Another function is suggested by FtsZ homology to eukaryotic tubulins. Like tubulin, FtsZ is a GTPase and undergoes GTP/GDP-dependent polymerization. Recent studies showed that Z-ring is a very dynamic structure suggesting that GTP-dependent

30

assembly/disassembly of Z-ring might provide constriction force to power cell division. Homologue of this gene identified in *Alloicoccus otitidis* is described in Example 5/Table 4 (Seq. ID No 83). The protein encoded by the gene is set forth in Seq. ID No. 84.

5

#### GTPase activity

FtsZ is a GTPase that contains the tubulin-signature nucleotide-binding motif GGGTGS/TG. Like in  $\alpha\beta$ -tubulin dimer, the active site for GTP-hydrolysis appears to be shared between two subunits where the GTP-binding pocket is provided by one subunit while the GTPase-activating T7 loop comes from the other subunit. This view is supported by genetic analysis as various mutations that inhibit FtsZ GTPase activity map in the T7-loop region and a conserved Asp-residue in T7-loop is found to be involved in the coordination of the cation involved in GTP hydrolysis. FtsZ GTPase activity is  $Mg^{2+}$ -dependent and is stimulated by KCl.

15

#### Polymerization

*In vivo*, about 75% of FtsZ is present as multimers. *In vitro*, FtsZ forms a variety of structures at various conditions. FtsZ assembles into thin protofilaments with GTP and formation of FtsZ polymers is coupled to GTP hydrolysis: when GTP runs out, polymers disassemble. Protofilaments assemble into sheets and bundles in the presence of multimolar amounts of either  $Mg^{2+}$  or  $Ca^{2+}$  or by addition of DEAE-dextran. In addition, ZipA protein induces bundling of FtsZ polymers. With GDP, FtsZ assembles into curved filaments and minirings.

#### 25 Interactions with other proteins

In *E. coli*, at least nine different proteins are localized to the division septum and are required for cell division to proceed. Among them two proteins, ZipA and FtsA, are shown to interact directly with FtsZ. Both of these proteins localize to the division site independently from each other, but require FtsZ for localization. ZipA is an integral membrane protein which is thought to mediate invagination of cell membrane by linking the membrane to constricting Z-ring. Interaction between ZipA and FtsZ is confined to C-terminal portion of ZipA (residues 185-328) and conserved 17-amino acid region on C-terminus of FtsZ. FtsA is an actin-like membrane-associated protein

30

which possesses ATPase activity and might provide energy required for Z-ring dynamics. Interaction between FtsZ and FtsA is not studied in great detail, it is shown that C-terminus of FtsZ is required. The remaining division proteins require both ZipA and FtsA for their localization to Z-ring.

5

### **FtsZ as a target for anti-infective development**

FtsZ is an essential protein for cell division/bacterial viability. Knock-out *ftsZ* mutants fail to divide and, as a result, filament and die. The target is widely conserved throughout bacterial kingdom implying that FtsZ-specific inhibitor would have a broad-spectrum antibacterial activity. The potential drawbacks of the target might include the presence and the essential role of a homolog (tubulin) in eukaryotes and an intrinsic difficulty in inhibiting protein-protein interactions by small molecules. Although this target is being studied extensively, no FtsZ-specific compounds are reported up to date.

15

### **Assays for measuring FtsZ function**

Polymerization of FtsZ is measured by light scattering assay as described previously. FtsZ (12.5  $\mu$ M) is incubated in 200  $\mu$ l of polymerization buffer (50 mM MES/NaOH, pH 6.5, 50 mM KCl, 5 mM MgCl<sub>2</sub>, 10 mM CaCl<sub>2</sub>) in a fluorescence cuvette with a 1 cm path length. The sample is maintained at 30°C, polymerization is induced by addition of 20-500  $\mu$ M GTP. Light scattering is measured at 90°, both excitation and emission wavelengths are set to 350 nm, slit width is 2 nm. Alternatively, the amount of polymerized FtsZ is analyzed by sedimentation and subsequent quantification of precipitated FtsZ by SDS-PAGE, Coomassie staining and densitometric scanning. In addition, polymers are observed by electron microscopy. This assay represents either primary or secondary/confirmatory assay.

25

GTP binding of FtsZ is monitored by the covalent cross-linking of [ $\gamma$ -<sup>32</sup>P]GTP (3000 Ci/mmol) to FtsZ in a previously described competition assay. FtsZ (3  $\mu$ g) is incubated in 20  $\mu$ l of 50 mM MES/NaOH, pH 6.5, 100 mM KCl, 4 mM MgCl<sub>2</sub>, 1 mM EDTA, 0.1 mM EGTA and 0.5 mM DTT. Various amounts of non-labeled competing nucleotide (GTP or GTP analogs) and 0.1 mM [ $\gamma$ -<sup>32</sup>P]GTP are added, samples are incubated at 0°C for 15 min, then UV cross-linked for 5 min and analyzed by SDS-

30

PAGE on 12% gel, autoradiography and densitometric scanning. This assay represents a secondary/confirmatory assay.

The GTP hydrolytic activity of FtsZ is monitored by thin-layer chromatography (TLC) as described previously. Briefly, the reaction mixture consists of 5 mM of [ $\gamma$ - $^{32}$ P]GTP (40 mCi/mmol), 15 mM magnesium acetate and 0.25-2 mg/ml of FtsZ in reaction buffer (40 mM Tris-acetate, pH7, 200 mM potassium acetate, 2 mM EDTA, 1 mM DTT and 0.5% Triton X-100), aliquots are separated by TLC and amount of GTP converted to GDP is determined by spot-densitometry. Alternatively, GTPase activity is measured either by quantitation of the non-radioactive inorganic phosphate with the malachite green-molybdate reagent as described previously or by quantitation by scintillation counting of radioactive inorganic phosphate released after hydrolysis of [ $\gamma$ - $^{32}$ P]GTP (26). This assay represents either primary or secondary/confirmatory assay.

Among interactions of FtsZ with various cell division proteins, interaction between FtsZ and ZipA is characterized the best. ZipA -induced bundling of FtsZ is measured by the light scattering assay that is described above, both proteins are used at  $\geq 5 \mu\text{M}$ .

#### EXAMPLE 22

#### 20 ALLOIOCOCCUS OTITIDIS ENCODED GYRA/GYRB (DNA GYRASE, TOPOISOMERASE II) AND GRGA/GRLB (TOPOISOMERASE IV)

DNA topoisomerases: topoisomerases modulate the topological state of DNA in cells. This involves binding to DNA, introducing single or double stranded breaks in the DNA, passing DNA molecules through the break and rejoining the break. This controls the levels of positive and negative supercoiling of DNA and functions in catenation/decatenation. Controlling the topological state of DNA is critical to the fundamental processes of transcription, recombination, replication and partitioning of the chromosome. There are two main categories of topoisomerases, type I and type II. Type I topoisomerases introduce single stranded breaks in DNA whereas type II enzymes introduce double stranded breaks. GyrA/GyrB (gyrase) and GrIA/GrIB (topoisomerase IV) are both type II enzymes that are essential for cell viability.

DNA gyrase (GyrA/GyrB) is a type II topoisomerase that functions to control the degree of supercoiling in double stranded DNA. It is essential for viability and



plays central roles in replication, repair, recombination and transcription of DNA.

Gyrases have the ability to introduce double stranded breaks in DNA molecules while remaining bound to the DNA through phosphotyrosine bonds, pass uncut DNA through the break and then rejoin the breaks, with repeated cycles being driven by

5 the hydrolysis of ATP. Gyrase has the unique ability to introduce negative supercoils in closed circular DNA and also functions to catenate/decatenate DNA duplexes.

The generation of negative supercoiling is important for initial stages in replication.

DNA gyrase from *Escherichia coli* has been studied in detail. It is a complex of two subunits of GyrA (encoded by *gyrA*) and two subunits of GyrB (encoded by *gyrB*) (ie.

10 A<sub>2</sub>B<sub>2</sub> complex). The subunits are organized in discrete domains. An N-terminal domain of GyrB harbors ATPase activity while the C-terminal domain is thought to interact with the GyrA subunit, and is involved in DNA binding. The N-terminal domain of GyrA is apparently involved in DNA strand breakage-ligation reactions while the C-terminal segment is involved in DNA binding. Crystal structures of the  
15 DNA strand breakage/reunion domain of *E. coli* GyrA, and the N-terminal ATPase domain of *E. coli* GyrB have been determined. DNA gyrase has also been purified and characterized from gram positive organisms such as *S. aureus*. Comparison of DNA gyrases from several bacteria reveal a high degree of conservation of important domains.

20 Topoisomerase IV (GrlA/GrlB) is a type II topoisomerase but unlike gyrase it does not possess negative supercoiling activity. Its primary role in replication appears to be in the decatenation of multiply linked daughter chromosomes, important for terminal stages of the replication process. Topoisomerase IV has been purified and characterized from gram negatives eg. *E. coli*, (where the GrlA/GrlB  
25 subunit homologs are designated ParC and ParE), and gram positives eg *S. aureus*. Homologs of these genes identified in *Alloicoccus otitidis* is described in Example 5/Table 4 (Seq. ID Nos 17 and 19). The proteins encoded by the genes are set forth in Seq. ID Nos. 18 and 20.

**GyrA/GyrB (Gyrase) and GrIA/GrIB (topoisomerase IV) as targets for anti-infective development:**

*Alloiococcus otitidis* is an infectious organism associated with disease, and consequently, novel antimicrobials to combat these infections are desirable. DNA gyrase and Topoisomerase IV is essential for bacterial viability and is a well-established and validated antibacterial target.

**Purification of DNA gyrase and topoisomerase IV from *Alloiococcus otitidis***

Genes encoding the GyrA/GyrB and GrIA/GrIB subunits or their functional domains are obtained using polymerase chain reaction amplification of the genomic region encoding them. The genes are then subcloned into standard expression vectors, with or without affinity tags. The enzyme is then overexpressed in *Escherichia coli* and purified using a standard tag system or conventional chromatography.

**Measurement of gyrase and topoisomerase IV by kinetoplast DNA decatenation assay:**

Type II topoisomerases introduce double stranded breaks in DNA and mediate catenation/decatenation of DNA. Topoisomerase IV activity is readily determined with decatenation assays using as substrate kinetoplast DNA (KDNA) from *Crithidia fasciculata*. The DNA isolated in this procedure is a highly networked series of catenated double stranded minicircles and is easily be pelleted by centrifugation. The activity of topoisomerase II enzymes results in the release of decatenated DNA minicircles from the networked KDNA. These have a high mobility in agarose gels and migrate into the gel ahead of the networked material, which has very low mobility, allowing for determination of decatenation activity using ethidium bromide stained agarose gel electrophoresis.

Alternatively, using radiolabeled KDNA, the level of decatenation activity is measured by counting radioactivity remaining in reaction supernatants following centrifugation to pellet the networked material. Typical conditions used for assaying decatenation activity of *S. aureus* and *E. coli* topoisomerase IV activity are as follows: *C. fasciculata* KDNA (0.9 mg/ml) is incubated in 40  $\mu$ l of reaction buffer (50 mM Tris-HCl, pH 7.7, 5 mM  $MgCl_2$ , 5 mM DTT, 50  $\mu$ g/ml bovine serum albumin, 1.5 mM ATP

and 350 mM potassium glutamate) with appropriate amounts of the Grl subunits, for 1 hour at 37° C. If non radiolabeled KDNA is used, these reactions can be stopped and analyzed by agarose gel electrophoresis, or for radioassays, the reaction is stopped by gentle mixing with 10 µl of stop solution (50 % glycerol, 50 mM EDTA (pH 8.0), 2.5 % SDS and 0.1 % bromophenyl blue) and centrifuged at 15 000 x g for 5 min at 20° C. Decatenation activity is determined by counting radioactivity in 25 µl of the supernatant in a scintillation counter. Alternatively, a modified assay employing flow injection fluorometry of 4', 6-diaminidino-2-phenylindole (DAPI) treated supernatants has been described that could be suitable for moderate throughput non radioactive assays, or filtration of the reactions through appropriate filters may efficiently separate the decatenated species from KDNA. Although the above described assays were used for topoisomerase IV, modified decatenation reactions using KDNA isolated from *Leishmania donovani* reveal significant decatenation activity by gyrase from *E. coli* and *Mycobacterium smegmatis*, indicating the applicability of the assay to prokaryotic gyrases.

#### DNA Supercoiling/relaxation assays.

DNA gyrase function is directly assayed using a simple supercoiling assay typified by that described for the measurement of *Escherichia coli* DNA gyrase activity. Briefly, incubation of relaxed closed circular plasmid DNA (pUC18, 7.5 nM) in the presence of DNA gyrase (approximately 10 nM) in 40 mM Tris-HCl (pH 8.0) buffer containing 25 mM KCl, 4 mM MgCl<sub>2</sub>, 2.5 mM spermidine and 1.4 mM ATP buffer results in the introduction of supercoils in the plasmid DNA. Changes in DNA supercoiling status are readily observed by the alteration of mobility of the DNA in agarose gels stained with ethidium bromide and comparison to the mobility of relaxed and supercoiled plasmid template. This strategy is employed for screening for DNA gyrase inhibitors.

Topoisomerase IV activity is assayed by measuring relaxation of supercoiled plasmid DNA. A typical relaxation assay used for *S. aureus* topoisomerase IV activity is as follows: topoisomerase IV enzyme and supercoiled plasmid DNA (pBR322, 0.6 µg) is incubated in 40 µl 50 mM Tris-HCl, pH 7.7, containing 5 mM MgCl<sub>2</sub>, 5 mM DTT, 50 µg/ml bovine serum albumin, 1.5 mM ATP, 5 mM spermidine and 20 mM KCl, for 30 min at 37°C. Changes in DNA supercoiling status can be

readily observed by the alteration of mobility of the DNA in agarose gels stained with ethidium bromide and comparison to the mobility of relaxed and supercoiled plasmid template

The ATPase activity of topoisomerases is measured using a coupled  
5 spectrophotometric ATPase assay described for the GyrB subunit of *E. coli*. ATPase activity is assayed in 300  $\mu$ l of 40 mM Tris-HCl (pH 8.0), containing 25 mM KCl, 2.5 mM spermidine, 4 mM  $MgCl_2$ , 400  $\mu$ M phosphoenolpyruvate, 250  $\mu$ M NADH, 3  $\mu$ l of pyruvate kinase /lactate dehydrogenase mix and ATP (0.5 – 3.5 mM). The reaction is started by the addition of truncated N-terminal derivatives of the GyrB protein (5  
10  $\mu$ M) containing the ATPase domain. ATPase activity is reflected as a decrease in absorbance of light at 340 nanometer wavelength.

#### DNA cleavage assay.

Quinolone drugs interfere with the DNA strand breakage-ligation cycle activity  
15 of many topoisomerases. Incubation of topoisomerase and linear or supercoiled pBR322 plasmid DNA, or small linear DNA fragments, in the presence of quinolones and magnesium results in the trapping of a complex of topoisomerase, DNA with a double stranded break and the drug. The topoisomerase remains bound to the cleaved DNA, however treatment with a denaturant such as SDS or proteinases  
20 remove/degrade the gyrase, releasing the cut DNA. Certain consensus sequences representing preferred cut sites of *E. coli* gyrase in plasmid pBR322 have been identified in template DNA molecules used in these assays. This assay is useful for mode of action studies of inhibitors of gyrase/topoisomerase IV activity and in particular of the strand breakage-ligation function. Cleavage reactions are performed  
25 with linear or supercoiled DNA. A typical cleavage reaction using linear DNA to measure cleavage by *E. coli* and *S. aureus* gyrase and topoisomerase IV in the presence of drugs is as follows: gyrase/ topoisomerase IV is incubated in 20  $\mu$ l 25 mM Tris-HCl (pH 7.5) containing 0.5 mM EDTA, 0.5 mM DTT, 3  $\mu$ g bovine serum albumin per ml, 10 mM  $MgCl_2$ , 120 mM KCL, 10 mM ATP, 10 000 dpm of 3' end  
30 labeled linear pBR322 plasmid DNA and drug for 1 hour at 37°C. (Note: for *S. aureus*, KCl is replaced with 0.7 M potassium glutamate). Reactions are terminated by adding 5  $\mu$ l 2.5% SDS-2.5 mg proteinase K per ml and incubating at 37°C for 30 minute, then adding 5  $\mu$ l 30% glycerol-1% SDS-50 mM EDTA-0.05 % bromophenol

blue. Cleavage products are resolved on 1% agarose gels and visualized by autoradiography.

Additional cleavage assays are also used that measure 1) the linearization of supercoiled plasmid DNA (pBR322), with linearization measured using scanning densitometry of DNA species separated on 1 % agarose gels, or 2) the cleavage of small linear DNA molecules of approximately 100 bp encompassing the preferred cleavage sequence 5'- GGCTGGATGGCCTTCCCCAT - 3' from position 990 in plasmid pBR322. In the latter case, the fragment is produced by PCR and radiolabeled with  $\gamma$ -<sup>32</sup>P ATP at the 5' end of the top strand. This DNA is incubated with 1.3 pmol DNA gyrase in a total volume of 10  $\mu$ l 35 mM Tris-HCl (pH 8.0), 24 mM KCl, 2 mM spermidine, 4 mM MgCl<sub>2</sub> and inhibitor compound at 37°C for 10 min. Reactions are stopped by addition of 8 mM EDTA and 1% SDS, then treated with 500  $\mu$ g/ml proteinase K for 2 hours at 37°C. The DNA is then cleaned by phenol-chloroform extraction and ethanol precipitation, resuspended in TE buffer (pH 8.0), and loaded and resolved on 12 % sequencing gels containing 7M urea. In the presence of inhibitors of the strand breakage-ligation function, radioactive cleavage products are detectable by autoradiography. Modifications of this assay whereby one strand of the DNA substrate is labeled with an affinity tag such as biotin and the other is radiolabeled or fluorescently labeled should facilitate rapid separation and detection of cleavage products using streptavidin coated columns or plates, resulting in higher assay throughput.

#### GYRASE ACTIVITY ASSAYS: DNA REPLICATION:

Early work by Fuller and Kornberg revealed that a partially purified crude soluble fraction derived from *Escherichia coli* cells (designated fraction II) contained the components necessary for replication of plasmids containing oriC (*E. coli* chromosomal origin of replication). Replication mediated by this fraction specifically required supercoiled plasmids. Although the exact makeup of the protein complex mediating the replication was not known, the replication reaction was inhibited by 1) rifampicin, and 2) nalidixic acid and novobiocin, indicating essential roles for both RNA polymerase and DNA gyrase, respectively. Subsequently the reaction was reproduced using replication machinery reconstituted from purified protein HU, DnaA,

DnaC, DnaB, single stranded binding protein (SSB), primase, DNA polymerase holoenzyme, RNA polymerase holoenzyme and GyrA/GyrB.

The requirement for gyrase activity for replication is exploited for the identification of gyrase inhibitors using a replication-based high throughput screen.

- 5 Gyrase specific inhibitors are identified from the overall pool of replication inhibitors using the secondary assays detailed below. Screening for inhibitors of gyrase in a setting where gyrase is participating in an overall reaction that is essential in bacteria might better select physiologically relevant inhibitors

- 10 An assay suitable for high throughput screening of inhibitors of replication (including gyrase and DnaA inhibitors) is based on the replication reaction of Kaguna and Kornberg. This reaction was set up as follows; standard reaction in 25  $\mu$ l: 40 mM Hepes (pH 7.6), 2 mM ATP, 0.5 mM GTP, CTP and UTP, 50  $\mu$ g/ml bovine serum albumin, 6 mM phospho creatine, 100 $\mu$ M dATP, dGTP, dCTP and dTTP,  $\gamma$ -<sup>33</sup>P dTTP (50-150 cpm/pmol total nucleotides) 11mM magnesium acetate, 100  $\mu$ g/mL creatine  
15 kinase, 85 ng SSB, 48 ng DnaB, 40 ng DnaC, 20 ng primase, 160 ng DNA polymerase III holoenzyme, 800 ng RNA polymerase, 150 ng GyrA, 350 ng GyrB, 120 ng DnaA, 2.5 units topoisomerase I, 190 ng HU, 0.15 ng Rnase H 200 ng supercoiled plasmid template. The reaction is assembled at 0 °C and initiated by incubation at 30°C. Replication reactions are terminated by the addition of EDTA to  
20 20 mM. Incorporation of nucleotides into DNA is measured by filtration through 96 well DEAE filter plates and counting retained radioactivity.

- Compounds inhibiting gyrase activity in *Alloiococcus otitidis* are found as part of a larger program directed at replication. This reaction described above uses the replication machinery of a gram-negative organism, which differs somewhat from the  
25 replication machinery of gram positives such as *Staphylococcus aureus* with respect to the specific protein subunits involved. Therefore a similar system specific to *Alloiococcus otitidis* is assembled from the relevant proteins purified from *Alloiococcus otitidis*. Several techniques are then utilized to determine the interaction of inhibitors with Gyr A and GyrB. These are described in the literature and include  
30 a) Nuclear magnetic resonance; and b) Capillary electrophoresis.

## Example 23

ALLOIOCOCCUS OTTIDIS ENCODED CELL WALL BIOSYNTHETIC ENZYMES MUR A

Bacterial cell wall peptidoglycan (murein) is a large macromolecule of periodic  
5 structure whose basic unit, a disaccharide-peptapeptide, is polymerized linearly via  
the disaccharide motif and cross-linked laterally via the peptide motif. The process of  
bacteria cell wall biosynthesis starts from the transferase MurA, which transfers the  
addition of an enolpyruvyl moiety to the 3'-hydroxyl-UDP-N-acetyl glycosamine  
(UDP-GluNAc). Subsequently, the reductase MurB reduces the enol ether to the  
10 lactyl ether, utilize one equiv. of NADPH and a solvent proton to form UDP-N-acetyl  
muramic acid (UDP-MurNAc). Next a series of ATP dependent amino acid ligases  
(MurC, MurD, MurE and MurF) catalyze the stepwise synthesis of the pentapeptide  
side chain using the newly synthesized carboxylate as the first acceptor site. Each  
enzyme is responsible for the addition of one more residue except MurF, catalyzes  
15 D-ala-D-ala. MurE in gram negative bacteria catalyzes the meso-2, 6-  
diaminopimelate (DAP), while in gram positive bacteria MurE catalyzes L-lysine.

The product of MurF, UDP-NAM pentapeptide is the final product of the  
cytoplasm enzymes and is the most important precursor for further peptidoglycan  
biosynthesis. UDP-MurNAc pentapeptide is then and catalyzed at the plasma  
20 membrane by the membrane bound enzymes such as the translocase MraY and  
transferase MurG.

UDP-N-acetylglucosamine enolpyruvyl transferase (MurA) catalyzes the first  
committed step in bacterial cell wall biosynthesis. The enzyme transfers an  
enolpyruvyl group from phosphoenolpyruvate (PEP) to UDP-N-acetylglucosamine  
25 (UDP-GluNAc) to the 3'-OH of UDP-GlcNAc by an addition-elimination mechanism  
that proceeds through a tetrahedral ketal intermediate. MurA product enolpyruvate  
UDP-N-acetylglucosamine (EP-UNAG) is a precursor to UDP- N-acetylmuramate  
(UDP-MurNAc), an essential building block for the bacterial cell wall. MurA is  
conserved across both gram-positive and gram-negative bacterial species: gram-  
30 negative bacteria have one copy of the *murA* and gram-positive bacteria have two  
copies. *Alloioococcus otitidis murA* was identified as described in Example 5/Table 4  
and its genomic structure set forth in Seq. ID No. 101. The amino acid sequence of  
the protein encoded by this gene is set out in Seq. Id No. 102.

**Alloiococcus otitidis murA as a target for anti-infective development**

*MurA* in *E. coli* and *Streptococcus pneumoniae* has been shown to be essential by gene deletion technique. The essentiality of *MurA* in gram-positive bacteria such as *Streptococcus pneumoniae* was demonstrated in that its deletion is fetal. No mammalian homolog to *MurA* has been reported. *MurA* is specifically inhibited by the natural product antibiotic fosfomycin. Thus the importance of *MurA* in peptidoglycan biosynthesis makes it an attractive target for the design of novel antibacterial agent.

**Assays for measuring MurA function****Phosphate detection:**

*MurA* activity is detected by quantitating the UDP-GluNAc-dependent Pi from PEP and assayed by Lanzetta's malachite Green-ammonium molybdate assay. Pi is quantitated by measuring the optical density at A660 nm.

**Coupled assay with MurB:**

A coupled assay in access of *MurB*, which reduces the *MurA* product EP-UNAG G to UDP-MurNAc, couples the *MurA* transferase activity with NADPH oxidation. The oxidation of NADPH is monitored at 340 nm and is stoichiometric with the production of EP-UNAG.

**Fluorescence experiments**

Fluorescence experiments to detect *murA* are performed using the hydrophobic fluorescence probe 8-anilino-1-naphthalene sulfonate (ANS). The fluorescence quenching of *MurA*/ANS solutions upon addition of UDP-GlcNAc or pyruvate-P is concentration dependent and in a saturating manner.

**Isothermal titration calorimetry**

The binding of UDP-GluNAc to *MurA* is studied in the absence and presence of the antibiotic fosfomycin by isothermal titration calorimetry. Fosfomycin binds covalently to *MurA* in the presence of UDP-GluNAc and also in its absence as



demonstrated by MALDI mass spectrometry. Novel Fosfomycin analogs and other antibiotics that bind to *murA* are also identifiable using isothermal titration chemistry.

#### Capillary electrophoresis-based enzyme assay

5           A capillary electrophoresis-based enzyme assay for MurA is described by Dai and colleagues. This method, based on UV detection, provides baseline separation of one of the reaction products, EP-UNAG, from substrates PEP and UDP-GlcNAc within 4 min. The other product, phosphate, is not detectable by UV at 200 nm. Quantitation of individual components, substrates or product, is be accomplished  
10       based on the separated peaks. This assay is also used to detect novel antibiotics, which inhibit *murA* activity.

#### EXAMPLE 23

##### ALLOIOCOCCUS OTITIDIS ENCODED CELL WALL BIOSYNTHETIC ENZYMES MURB

15

MurB, the UDP-*N*-acetyl enolpyruvyl glucosamine reductase, commits the second step of bacterial cell wall biosynthesis in cytoplasm and is responsible for the reduction of the enol ether to the lactyl ether, utilizes one equiv. of NADPH and a solvent proton. The product of MurB is UDP-*N*-acetylmuramic acid (UDP-MurNAc), the linker of the peptide  
20       and glycan portions of cell wall precursor UDP muramyl-pentapeptide. MurB from *E. coli* is a 342 amino acid protein, which has a distinctive yellow color characteristic of bound flavin as its co-factor. The biochemistry characterization and X-ray crystal structure of MurB in *E. coli*, in *Staphylococcus aureus* and *Streptococcus pneumoniae* have been studied extensively. The gene *Alloiococcus oitidis murB* was identified as disclosed as  
25       described in Example 5, and is set out in Seq. ID No. 39. The amino acid sequence of the protein encoded by this gene is set out in Seq. ID No. 40.

##### ***Alloiococcus oitidis murB* as a target for anti-infective development**

30           The essentiality and unique function of MurB in prokaryotic cells and the absence of homologue in eukaryotic cells make it an attractive novel antibacterial target. To date, no small molecule inhibitors of MurB have been reported.

*Alloicoccus oititidis* ORF-1263 (*murB*) (Seq. ID No. 39) encodes enzyme UDP-N-acetylenolpyruvylglucosamine Reductase (MurB) as shown by sequence homology.

5     **Assays for measuring *MurB* activity**

**Spectrophotometric assay monitoring NADPH consumption:**

      MurB activity is typically monitored by its biochemical reaction in which NADPH reduces the bound FAD and resulting decrease in absorbance at 340 nm. Enzyme is maximally activated in the presence of K<sup>+</sup>, NH<sup>4</sup> at cation concentrations  
10    between 10-50 mM.

**Coupled assay with *MurC*:**

      In designing an end point assay for high through put screen (HTS), a novel coupled assay in access of UDP-MurNAc L-alanine synthase (*MurC*) was developed  
15    at Wyeth. This assay utilizes the biochemically synthesized *MurA* product EP-UNAG as substrate, coupled with limited *MurB* and excess *MurC* in the reaction with all other substrates/components involved. In this assay, *MurB* is responsible for the reduction of the enol ether to the lactyl ether, and the follow up enzyme *MurC* catalyzes the ATP dependent ligation of the first of the five amino acids of UDP-  
20    peptide with a release of one molecule of phosphate. After 60 minutes of incubation, color reagent malachite green was added and phosphate was detected spectrophotometrically.

**Fluorescence binding assay**

25       A fluorescence method developed at Wyeth is used to determine the binding potency (K<sub>d</sub> value), stoichiometry and nature of binding site of substrates and inhibitors interactions with *MurB* enzymes. This assay is based on changes in intrinsic fluorescence of inhibitor and/or enzyme, upon formation of enzyme-inhibitor complex. Oxidized form of *MurB* consists of two fluorescent groups, namely  
30    tryptophan residues and the cofactor FAD. Upon binding inhibitor or substrate, local changes in the solvent environment of these groups or overall conformational and electronic changes occur in the enzyme due to which the fluorescence emission is altered. For instance, inhibitor binding significantly quenched the fluorescence and

altered the solvent environment of FAD to a less polar environment. The changes in the fluorescence of the FAD moiety are used to estimate binding constants for *MurB* inhibitors. Binding experiments are set up in which a fixed concentration of enzyme is titrated with increasing concentrations of the inhibitor. In typical inhibitor binding experiments, the fluorescence emission of the FAD moiety is quenched due to specific interactions of the inhibitor with *MurB* enzymes and the binding site was saturated at micromolar concentrations of inhibitor. The changes in the fluorescence are fitted to mathematical binding models to determine binding affinity.

#### 10 Temperature-jump isothermal denaturation procedure

Temperature-jump isothermal denaturation procedure with various methods of detection is used to evaluate the quality of putative inhibitors of *MurB* discovered by high-throughput screening. Three optical methods of detection-ultraviolet hyperchromicity of absorbance, fluorescence of bound dyes, and circular dichroism-as well as differential scanning calorimetry are used to dissect the effects of two chemical compounds and a natural substrate on the enzyme. The kinetics of the denaturation process and binding of the compounds detected by quenching of flavin fluorescence are used to quantitate the dose dependencies of the ligand effects.

#### 20 NMR studies

NMR studies are performed using perdeuterated, uniformly  $^{13}\text{C}/^{15}\text{N}$ -labeled samples of *MurB*. In the case of substrate-free *MurB*, one or more backbone atoms are assigned for 334 residues (96%). For NADP $^{+}$ -complexed *MurB*, one or more backbone atoms are assigned for 313 residues. The strategies used for obtaining resonance assignments are known. Localizing the NADP $^{+}$  binding site on the *MurB* enzyme is also studied by NMR methodology.

### EXAMPLE 25

#### ALLOIOCOCCUS OTITIDIS ENCODED CELL WALL BIOSYNTHETIC ENZYME, *MURC*

30

Uridine diphosphate-N-acetylmuramate:L-alanine ligase (*MurC*) catalyzes the third chemical step of bacterial cell wall biosynthesis. This enzyme is a nonribosomal peptide ligase which utilize ATP to form an amide bond between L-alanine and UDP-

N-acetylmuramic acid (UDP-MurNAc). This ATP-dependent ligation adds the first of five amino acids to the sugar moiety of the peptidoglycan precursor. Also, in this reaction, ATP is converted to ADP with release of one molecule of inorganic phosphate. Thus MurC reaction is an essential step in cell wall biosynthesis for both gram-positive and gram-negative bacteria. The genetic, biochemistry analysis and crystal graphic studies of MurC in gram-negative bacteria *E. coli* have been extensively studied. Characterizations of MurC in other pathogens such as *Staphylococcus aureus* and *Pseudomonas aeruginosa* have also been documented.

10 ***Alloiococcus otitidis* encoded MurC as a target for anti-infective development**

The *Alloiococcus otitidis* ORF-2602 (*murC*, Seq. ID No. 95) encodes enzyme UDP-MurNAc:L-alanine ligase (*MurC*) as determined by sequence homology. This enzyme presents a target for the development of novel anti-infectives to treat the disease(s) caused by this pathogen. Novel compounds identified using combinatorial chemistries are assayed for their inhibitory effect on *MurC* activity using one of the assays set out below.

**Assays for measuring MurC activity**

20 **Spectrophotometric assay detecting phosphate release:**

MurC activity is detected by the inorganic phosphate production. Typically the reaction mixture contains substrates ATP, L-alanine, UDP-MurNAc, DTT, MgCl<sub>2</sub> and MurC enzyme. After 20 minutes incubation, the reaction is quenched with the addition of malachite Green-ammonium molybdate for a colored reaction. Absorbance at 660 nm is read 5 minutes after the quench. Absorbance values are converted to concentration of Pi with standard curves using KH<sub>2</sub>PO<sub>4</sub>, which is prepared under identical conditions without the enzyme MurC.

**Spectrophotometric assay detecting formation of ADP**

30 Due to the conversion of ATP to ADP in MurC reaction, the production of ADP is monitored in coupled enzymes spectrophotometrically. In this reaction, in addition to MurC substrate UDP-MurNAc, L-alanine and ATP, NADH, phosphoenolpyruvate, MgCl<sub>2</sub> and (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, two other coupled enzymes pyruvate

kinase and lactase dehydrogenase are also presented. Reaction mixtures without ATP and MurC are incubated at 37°C for 10 min before ATP is added for another minute. Reaction is then started by the addition of MurC. The decrease of NADH absorbance at 340 nm is monitored spectrophotometrically. One unit of activity  
5 corresponds to 1  $\mu$ mol of ADP formed per hour.

#### L-Alanine radio-labeled assay:

The MurC enzyme activity in this assay is measured as endpoint using  $^{14}\text{C}$ -L-alanine and ATP incubated with  $\text{MgCl}_2$ , and  $(\text{NH}_4)_2\text{SO}_4$  in 100 mM Tris/HCl, pH 8.0.  
10 Reaction is initiated by the addition of the catalytic amounts of MurC. Samples of the reaction mixture are then mixed with glacial acetic acid and then stored at 4°C. Remaining  $^{14}\text{C}$  -L-alanine is separated from  $^{14}\text{C}$  -UDPMurNAc on SCX columns run under vacuum. Quenched reaction samples are supplemented with equilibration buffer and counted using a liquid scintillation counter.

15

#### EXAMPLE 26

##### ALLOIOCOCCUS OTTIDIS ENCODED CELL WALL BIOSYNTHETIC ENZYMES MURD

Bacterial UDP-N-acetylmuramyl-L-alanine:D-glutamate ligase (MurD), a  
20 cytoplasmic peptidoglycan biosynthetic enzyme, catalyzes the fourth step of bacterial cell wall biosynthesis. In this reaction, MurD catalyzes ATP-dependent addition of D-glutamate to an alanyl residue of the UDP-N-acetylmuramyl-L-alanine (UDP-MurNAc-L-Ala) precursor, generating the UDP-MurNAc-dipeptide. The formation of a peptide linkage between the amino function of D-glutamate and the carboxy  
25 terminus of UDP-N-acetylmuramyl-L-alanine is generated through this reaction. The stoichiometric consumption of ATP supplies the energy needed for this peptide bond formation with concomitant generation of ADP and orthophosphate. The *murD* genes were cloned and characterized from gram-positive bacteria of *Staphylococcus aureus* and *Streptococcus pyogenes*, and gram-negative bacteria from *Escherichia coli*, *Haemophilus influenzae*, *Bacillus subtilis*. Structures of MurD from *E. coli* and  
30 MurD complexed with its substrate UDP-MurNAc-L-Ala have been solved to 2.0 Å resolution. The role of specific amino acids at the active site of MurD have been extensively studied using the ortholog and paralog amino acid invariants. Homologue

of this gene identified in *Alloiococcus otitidis* is described in Example 5/Table 4 (Seq. ID No 89). The protein encoded by the gene is set forth in Seq. ID No. 90.

#### **Alloiococcus otitidis encoded MurD as a target for anti-infective development**

5

Due to its high specificity and essentiality, MurD is an attractive target for the development of novel antimicrobial agents. *Alloiococcus otitidis* ORF-2494, by sequence homology, has been shown to encode enzyme UDP-N-acetylmuramyl-L-alanine:D-glutamate ligase (MurD) (Seq. ID. No. 89). Inhibition of MurD activity is used to identify novel antimicrobial agents.

10

#### **Assays for measuring MurD activity**

##### **Spectrophotometric assay detecting phosphate release:**

15

MurD activity in the presence or absence of a putative inhibitory molecule of MurD is detected by the orthophosphate production in test tube or in 96-well format. Typically the reaction mixture contains substrates ATP, D-glutamine, UDP-MurNAc-L-Ala, DTT, MgCl<sub>2</sub> and MurD enzyme. After 20 minutes incubation, the reaction is quenched with the addition of malachite Green-ammonium molybdate for a colored reaction. Absorbance at 660 nm is read 5 minutes after the quench using Molecular Devices SpectraMax 250 plate reader. Absorbance values are converted to concentration of Pi using orthophosphate standards, which are prepared under identical conditions without the enzyme MurD.

20

25

##### **Spectrophotometric assay for detecting formation of ADP in the presence or absence of a putative inhibitory molecule of MurD:**

Due to the conversion of ATP to ADP in MurD reaction, the production of ADP is monitored with coupled enzymes of pyruvate kinase and lactate dehydrogenase spectrophotometrically. In this reaction, in addition to MurD substrate UDP-MurNAc-L-ala and ATP, MgCl<sub>2</sub> and (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, there is also in significant access of NADH, phosphoenolpyruvate, and two coupled enzymes pyruvate kinase and lactate dehydrogenase. This protocol monitors ADP formation

30

in the MurD catalyzed reaction, in the presence or absence of a putative inhibitory molecule of MurD, by the decrease of NADH absorbance at 340 nm.

**L-Glutamate radio-labeled assay:**

5 The MurD enzyme activity in the presence or absence of putative inhibitors of MurD is also measurable using D-<sup>14</sup>C- glutamate as an endpoint assay. The reaction mixture contains D-<sup>14</sup>C- glutamate UDP-MurNAc-L-Ala, ATP, MgCl<sub>2</sub>, (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> in 100 mM Tris/HCl, pH 8.0. An HPLC assay with online UV and flow scintillation detects the formation of UDP-MurNAc-L-Ala-D-<sup>14</sup>C Glu and ADP in each reaction.

10 **EXAMPLE 27**

**ALLOIOCOCCUS OTITIDIS ENCODED CELL WALL BIOSYNTHETIC ENZYME, MUR E**

15 The fifth step in the cytoplasmic peptidoglycan biosynthetic is catalyzed by MurE. In this step, the monomer units in the *Escherichia coli* and *Staphylococcus aureus* cell wall peptidoglycans differ in the nature of the third amino acid in the L-alanyl-gamma-D-glutamyl-X-D-alanyl-D-alanine side chain, where X is meso-diaminopimelic acid or L-lysine, respectively. Therefore, MurE from *E. coli* is the UDP-N-acetylmuramoyl-L-alanyl-D-glutamate: meso-diaminopimelic acid ligase, and 20 MurE from *S. aureus* is the UDP-N-acetylmuramoyl-L-alanyl-D-glutamate: L-lysine ligase. Thus represents the major difference of MurE from other murein enzymes in cytoplasm. The amino acid residues catalyzed by MurE plays a key role in the integrity of sacculus since it is directly involved in the peptide cross-linkage. MurE reaction is also ATP-dependent, which supplies the energy needed for the peptide 25 bond formation with concomitant generation of ADP and orthophosphate.

The essentiality of *MurE* has been well documented in *E. coli*, in *S. aureus*, as well as other pathogens such as *Haemophilis influenzae*, *Vibrio cholerae* and *Corynebacterium glutamicum*. Gene *murE* has been shown to be essential in bacteria. Homologue of this gene identified in *Alloiococcus otitidis* is described in 30 Example 5/Table 4 (Seq. ID No 25). The protein encoded by the gene is set forth in Seq. ID No. 26.

**Alloiococcus otitidis MurE as a target for anti-infective development**

*Alloiococcus otitidis* ORF-851, by sequence homology encodes enzyme UDP-N-acetylmuramyl-L-alanine-D-glutamate ligase: meso-diaminopimelic acid/or L-Lysine (MurE) (Seq. ID No 25). MurE activity in the presence or absence of a putative inhibitory molecule of MurE activity is used to identify novel antimicrobial agents, which may be used to treat disease caused by *Alloiococcus otitidis*.

#### Assays for measuring MurE activity

##### Radio labeled substrate assay: meso-A2pm-adding activity

Activity of MurE from *Alloiococcus otitidis* in the presence or absence of a putative inhibitory molecule of MurE activity is measured by using radio-labeled meso-<sup>14</sup>C A2pm mixing with ATP, MgCl<sub>2</sub>, UDP-MurNAc-L-Ala-D-Glu, DTT in 100 mM Tris/HCl and MurE from *Alloiococcus otitidis*.

##### Radio labeled substrate assay: L-lysine adding activity

Activity of MurE from *Alloiococcus otitidis* in the presence or absence of a putative inhibitory molecule of MurE activity is measured by using radio-labeled UDP-MurNAc-L-Ala-D-<sup>14</sup>C-Glu mixing with ATP, MgCl<sub>2</sub>, DTT, L-lysine in 100 mM Tris/HCl and MurE from *Alloiococcus otitidis*.

In both cases, mixtures are incubated at 37°C for 30 min, and reactions stopped by the addition of acetic acid. Reaction product is separated by high voltage electrophoresis in 2% formic acid for 45 min. The radio active spots corresponding to substrate and reaction product are detected by overnight autoradiography, or with radio scanner. The spots are also cut out and counted using liquid scintillation counter.

#### Example 28

##### ALLOIOCOCCUS OTITIDIS ENCODED CELL WALL BIOSYNTHETIC ENZYME, MURF

The D-alanyl-D-alanine-adding enzyme MurF encoded by the *murF* gene catalyzes is the last step of the cytoplasmic peptidoglycan biosynthesis. MurF performs the ATP-dependent formation of UDP-N-acetylmuramyl-L-gamma-D-Glu-meso-diaminopimelyl-D-Ala-D-Ala (UDP-MurNAc-pentapeptide). The product of MurF, UDP-MurNAc pentapeptide, is the final product of the cytoplasm enzymes and



is the most important precursor for further peptidoglycan biosynthesis. UDP-MurNAc  
pendapeptide is then catalyzed by the plasma membrane bound enzymes such as  
the translocase MraY and transferase MurG. Homologue of this gene identified in  
*Alloiococcus otitidis* is described in Example 5/Table 4 (Seq. ID No 3). The protein  
5 encoded by the gene is set forth in Seq. ID No. 4.

#### ***Alloiococcus otitidis* MurF as a target for anti-infective development**

Due to its high specificity, essentiality, and importance of its product UDP-  
MurNAc pentapeptide, MurF is attractive as an antibacterial target. The *Alloiococcus*  
10 *otitidis* ORF-48, by sequence homology, encodes enzyme UDP-N-acetylmuramyl-L-  
alanine-D-glutamate ligase: meso-diaminopimelic acid/or L-Lysine -alanyl-D-alanine-  
adding enzyme (MurF) (Seq. ID No. 3). MurF activity in the presence or absence of a  
putative inhibitory molecule of MurF activity is used to identify novel antimicrobial  
agents, which may be used to treat disease caused by *Alloiococcus otitidis*.

#### **Assays for measuring MurF activity**

##### **Spectrophotometric assay detecting phosphate release:**

Activity of MurF from *Alloiococcus otitidis* in the presence or absence of a  
20 putative inhibitory molecule of MurF activity is detected by the inorganic phosphate  
release in the ATP dependent MurF reaction. This assay detects nonomole amount  
of Pi in the reaction mixture contains substrates ATP, D-ala-D-ala, UDP-MurNAc-  
tripeptide, DTT, MgCl<sub>2</sub> and MurF enzyme. After 5 minutes incubation, the reaction is  
quenched with the addition of malachite Green-ammonium molybdate for a colored  
25 reaction.

##### **Coupled spectrophotometric assay detecting formation of ADP**

Due to the conversion of ATP to ADP in MurF reaction, the production of ADP  
in the presence or absence of a putative inhibitory molecule of MurF activity, is  
30 monitored with coupled enzymes of pyruvate kinase and lactate dehydrogenase  
spectrophotometrically. In this reaction, the decrease at 340 nm is observed as  
NADP is consumed in MurF reaction process. The reaction typically contains tris

buffer, substrates ATP, D-ala-D-ala, UDP-MurNAc-tripeptide, DTT, MgCl<sub>2</sub>, phosphoenopyruvate, NADPH and MurF enzyme.

#### EXAMPLE 29

##### 5 **ALLOIOCOCCUS OTITIDIS ENCODED CELL WALL BIOSYNTHETIC ENZYME, MURG**

MurG, the last enzyme involved in the intracellular phase of peptidoglycan synthesis, is a membrane-associated glycosyltransferase. MurG catalyzes the transfer of *N*-acetyl glucosamine from UDP to the C4 hydroxyl of a lipid-linked *N*-  
10 acetyl muramic acid derivative (lipid I) to form lipid II. Lipid II is a linked disaccharide that is the minimal subunit of peptidoglycan. Once lipid II is formed, this disaccharide is translocated across the bacterial membrane where it is polymerized and cross-linked to form the peptidoglycan layers. MurG has been shown to be essential for bacterial survival. The inactivation of MurG gene rapidly inhibits peptidoglycan  
15 synthesis in exponential growing cells. As a result, various alterations of cell shape are observed, and cell lysis finally occurs. Homologue of this gene identified in *Alloiococcus otitidis* is described in Example 5/Table 4 (Seq. ID No 87). The protein encoded by the gene is set forth in Seq. ID No. 88.

##### 20 ***Alloiococcus otitidis* MurG as a target for anti-infective development**

MurG is shown to be associated with the inner face of cytoplasmic membrane, and establishing that the entire peptidoglycan monomer unit assembled before being transferred across the membrane. MurG is a key enzyme at the border  
25 line between cytoplasmic and membrane of peptidoglycan synthesis, thus makes it an attractive target for novel antibacterial agent. Further, no mammalian analogues of MurG have been identified. Due to its high specificity, essentiality, and importance, MurG is attractive as an antibacterial target.

The *Alloiococcus otitidis* ORF-2492 has been shown to encode, by sequence  
30 homology, glycosyltransferase (MurG) (Seq. ID No.....). MurG activity in the presence or absence of a putative inhibitory molecule of MurG activity is used to identify novel antimicrobial agents, which may be used to treat disease caused by *Alloiococcus otitidis*.

## Assays for measuring MurG function

### Radiolabeled reaction

Activity of MurG from *Alloiococcus otitidis* in the presence or absence of a putative inhibitory molecule of MurG activity is measured by using  $^{14}\text{C}$  labeled N-UDP-GluNAc in the reaction containing UDP-MurNAc-pentapeptide,  $\text{MgCl}_2$ , ATP and MurG protein. The reaction is stopped after 30 min incubation and by boiling for 3 min. The reaction mixtures are applied to a Whatman I filter paper and subject to descending chromatography overnight. Radioactivity is located and countered with a scanner. This assay is also used to identify the specificity of inhibitor of MraY or MurG, based on the detection of radiolabeled  $^{14}\text{C}$  GluNAc incorporated into membrane precursors.

### Fluorometric assay

Based on the decrease in NADPH fluorescence at 465 nm, MurG reaction is also monitored in a reaction mixture of HEPES buffer,  $\text{MgCl}_2$ , Triton, phosphoenolpyruvate, and coupled enzymes of lactic dehydrogenase and pyruvate kinase, UDP-GluNAc and synthesized lipid I analogue in the presence or absence of putative inhibitors of MurG activity. One micromolar UDP corresponds to 500-fluorescence unit under the instrument setting.

## EXAMPLE 30

### ALLOIOCOCCUS OTTIDIS ENCODED BY HMG CoA REDUCTASE (MVA A)

Two pathways for isopentenyl diphosphate (IPP) synthesis have been described in bacteria: the mevalonate pathway and the non-mevalonate (MEP or GAP-pyruvate) pathway. The mevalonate pathway predominates in the archaeobacteria, gram-positive organisms, yeast and mammals; whereas the MEP pathway is found in gram-negative organisms, *B. subtilis*, chlamydia, and mycobacterium. The first HMG CoA reductase gene to be sequenced was cloned from *P. mevalonii*, in which HMG CoA reductase permits growth on mevalonate as a sole carbon source. A number of genes of the mevalonate pathway were identified in *S. aureus*, *S. epidermidis*, *S. pyogenes*, *S. pneumoniae*, *E. faecalis* and *E. faecium*. One of the genes, which encodes for HMG-CoA reductase (*mvaA*), when deleted

severely attenuated for virulence in a mouse model indicating that *mvaA* is essential. Due to its high specificity, essentiality, and importance, *mvaA* is attractive as an antibacterial target. Homologue of this gene identified in *Alloicoccus otitidis* is described in Example 5/Table 4 (Seq. ID No 37). The protein encoded by the gene is set forth in Seq. ID No. 38.

#### HMG-CoA reductase (MvaA) as a target for anti-infective development

The *Alloicoccus otitidis* ORF- has been shown to encode, by sequence homology, HMG-CoA reductase (*mvaA*) (Seq. ID No 37). *MvaA* activity in the presence or absence of a putative inhibitory molecule of HMG-CoA reductase (*mvaA*) activity is used to identify novel antimicrobial agents, which may be used to treat disease caused by *Alloicoccus otitidis*.

#### Assays for measuring HMG-CoA reductase (*mvaA*) activity

*MvaA* is purified by standard methods using widely available molecular tags following expression at high level from *E. coli*. Enzymatic activity is monitored in the presence or absence of a putative inhibitory molecule of HMG-CoA reductase activity by following oxidation of NADPH to NADP spectrophotometrically at 340 nm. The assay is carried out in the following buffer: 0.25 mM NADPH, 0.25 mM HMG-CoA, 50 mM NaCl, 1 mM EDTA, 5 mM DTT, 25 mM  $\text{KH}_2\text{PO}_4$  (pH 7.5). The assay is amenable to HTS in high density screening microtiter plates.

**Forward reaction:** Activity of HMG-CoA reductase (*mvaA*) from *Alloicoccus otitidis* in the presence or absence of a putative inhibitory molecule of HMG-CoA reductase activity is measured by reductive deacylation of HMG-CoA to mevalonate as measured the consumption of NADPH to NADP. Unlike other class II HMG Coa reductases, *MvaA* from *Alloicoccus otitidis*, like *S. aureus*, can use either NADPH or NADH cofactor in the reaction. The following kinetic data describe the reaction:  $K_m(\text{HMG CoA}) = 40 \mu\text{M}$ ,  $K_m(\text{NADPH}) = 70 \mu\text{M}$ ,  $K_m(\text{NADP}) = 100 \mu\text{M}$  (12). This assay is inhibitable by the statin drug fluvastatin; the  $K_i$  was measured at  $320 \mu\text{M}$ , which is four orders of magnitude higher than the  $K_i$  for class I HMG-CoA reductases.

**Reverse reaction:** The oxidative acylation of mevalonate to HMG-CoA in the presence or absence of a putative inhibitory molecule of HMG-CoA reductase activity is also monitored. The following kinetic data describes the reaction:  $K_{m(\text{mevalonate})} = 670 \mu\text{M}$ ,  $K_{m(\text{CoASH})} = 390 \mu\text{M}$ ,  $K_{m(\text{NADP})} = 580 \mu\text{M}$  (12).

5

### EXAMPLE 31

#### ALLOIOCOCCUS OTITIDIS ENCODED DIPHOSPHOMEVALONATE DECARBOXYLASE (MvAD)

Diphosphomevalonate decarboxylase, encoded by *mvaD*, the final enzyme acting in the mevalonate pathway of IPP synthesis was cloned from *S. aureus* by Wilding *et al* in 2000. Insertional inactivation of *mvaD* could only be accomplished when the strains were supplemented with mevalonate, indicating that *mvaD* is essential. The final step of the mevalonate pathway leading to IPP is the decarboxylation and dehydration of mevalonate-5-pyrophosphate to form isopentenyl diphosphate by MvaD (diphosphomevalonate decarboxylase).

MvaD homologues are well represented in gram-positive organisms (10). Phylogenetic analysis revealed that the cluster of gram-positive enzymes (39-80% identity) were well separated from the eukaryotic homologues, suggesting utility as an antibacterial target. The *Alloiococcus otitidis* ORF- 1275b has been shown to encode, by sequence homology, diphosphomevalonate decarboxylase (MvaD) (Seq. ID No. 43). MvaD activity in the presence or absence of a putative inhibitory molecule of diphosphomevalonate decarboxylase (MvaD) activity is used to identify novel antimicrobial agents, which may be used to treat the disease(s) caused by *Alloiococcus otitidis*. The protein encoded by the gene is set forth in Seq. ID No. 44.

25

## Example 32

ALLOIOCOCCUS OTITIDIS ENCODED HMG CoA SYNTHASE (MVAS)

The second step of the mevalonate pathway leading to IPP is the irreversible  
5 condensation of acetoacetyl-CoA and acetyl-CoA to form HMG-CoA by MvaS (HMG  
CoA synthase). It has been shown that *mvaS* knockout mutant of *S. pneumoniae*  
was attenuated for virulence. Due to its high specificity, essentiality, and importance,  
*mvaS* is attractive as an antibacterial target. Homologue of this gene identified in  
*Alloiooccus otitidis* is described in Example 5/Table 4 (Seq. ID No 35). The protein  
10 encoded by the gene is set forth in Seq. ID No. 36.

**HMG CoA SYNTHASE (MVAS) AS A TARGET FOR ANTI-INFECTIVE DEVELOPMENT**

The *Alloiooccus otitidis* ORF- has been shown to encode, by sequence  
15 homology, MvaS (HMG CoA synthase) (Seq. ID No. 35). MvaS activity in the  
presence or absence of a putative inhibitory molecule of HMG-CoA synthase (*mvaS*)  
activity is used to identify novel antimicrobial agents, which may be used to treat  
disease caused by *Alloiooccus otitidis*.

**Assays for measuring MvaS function**

MvaS is purified by standard methods using widely available molecular tags  
following expression at high level from *E. coli*. HMG-CoA synthase activity in the  
presence or absence of a putative inhibitory molecule of HMG-CoA synthase (*mvaS*)  
is assayed by measuring the loss of the enolate form of acetoacetyl-CoA  
25 spectrophotometrically. The reaction is carried out in a buffer containing 50 mM Tris  
(pH 9.75), 5.0 mM MgCl<sub>2</sub>, 500 μM acetyl-CoA, 20 μM acetoacetyl-CoA and enzyme.  
The enolate formed is monitored at 302 nm; therefore, as the acetoacetyl-CoA is  
consumed the signal is depleted. Using this assay the following kinetic data is  
measured:  $K_{m(\text{acetyl-CoA})} = 350 \mu\text{M}$ ;  $K_m^{\text{app}}(\text{acetoacetyl-CoA}) = 10 \mu\text{M}$ . This assay is amenable  
30 to HTS in high- high density screening microtiter plates.

**Example 33****ALLOIOCOCCUS OTITIDIS ENCODED NICOTINAMIDE ADENINE DINUCLEOTIDE ADENYLYL  
TRANSFERASE (NADD)**

5 Nicotinamide adenine dinucleotide (NAD) is an essential molecule in all living  
cells. NAD is synthesized via a multi-step *de novo* pathway or via a pyridine salvage  
pathway. The enzyme nicotinic acid mononucleotide adenylyl transferase (NaMN AT,  
EC2.7.7.18) catalyzes the conversion of ATP and nicotinic acid mononucleotide  
(NaMN) to nicotinic acid adenine dinucleotide (NaAD). The *nadD* gene, encoding  
10 bacterial NaMN AT, is essential for NAD biosynthesis and bacterial cell survival.  
NadD contains well-conserved the nucleotidyl transferase consensus sequence  
(GXFXXHXGH). The adenylyl transferase encoded by the *nadD* gene prefers  
NaMN over nicotinomide mononucleotide (NMN) as substrate. Due to its high  
specificity, essentiality, and importance, *nadD* is attractive as an antibacterial target.  
15 Homologue of this gene identified in *Alloiococcus otitidis* is described in Example  
5/Table 4 (Seq. ID No 91). The protein encoded by the gene is set forth in Seq. ID  
No. 92.

20 **NICOTINAMIDE ADENINE DINUCLEOTIDE ADENYLYL TRANSFERASE (NADD)  
AS A TARGET FOR ANTI-INFECTIVE DEVELOPMENT**

The *Alloiococcus otitidis* ORF- has been shown to encode, by sequence  
homology, nicotinomide adenine dinucleotide adenylyl transferase (NadD) (Seq. ID No.  
91). NadD activity in the presence or absence of a putative inhibitory molecule of  
25 NadD activity is used to identify novel antimicrobial agents, which may be used to  
treat disease caused by *Alloiococcus otitidis*.

**Assays for measuring NadD function****Discontinuous assay**

30 NadD activity in *Alloiococcus otitidis* is measured in the presence or  
absence of a putative inhibitory molecule of NadD activity. NadD converts  
nicotinic acid mononucleotide (NaMN) and adenosine triphosphate (ATP) to  
nicotinic acid dinucleotide (NaAD) and pyrophosphate (PP<sub>i</sub>). Each PP<sub>i</sub>  
molecule produced by the NadD reaction is then converted to two phosphate

(P<sub>i</sub>) molecules in the presence of inorganic pyrophosphatase (PPase). The P<sub>i</sub> molecules present are quantitated with a malachite green reagent at 660 nm.

**HPLC-based assay:** Enzyme activity is measured by HPLC quantitation of the reaction products. A neutralized aliquots from the reaction described above was injected into an HPLC system utilizing a 250 x4.6 mm Supelcosil LC-18 5µm reversed-phase column. The elution conditions: 9 min at 100% buffer A (0.1 M potassium phosphate buffer, pH6.0, 6 min at up to 12% buffer B (buffer a, containing 20% methanol, 2.5 min at up to 45% buffer B, 2.5 min at up to 100% buffer B, and hold at 100% buffer B for 5.5 min. The eluate absorbance was monitored at 254 nm.

#### Continuous assay

In bacteria, NadD combines nicotinic acid mononucleotide (NaMN) and adenosine triphosphate (ATP) to form nicotinic acid adenine dinucleotide (NaAD). NadE then converts NaAD into nicotinamide adenine dinucleotide (NAD) in the presence of ammonia and ATP. In the assay, the NAD product is reduced to NADH with alcohol dehydrogenase (ADH) and ethanol, thus permitting direct spectrometric detection of NADH at 340 nm wavelength. The coupled reaction above also includes inorganic pyrophosphatase (PPase) to prevent accumulation of the pyrophosphate byproduct from the consumption of ATP.

#### EXAMPLE 34

##### ALLOIOCOCCUS OTTIDIS ENCODED NICOTINAMIDE ADENINE DINUCLEOTIDE SYNTHASE (NADE)

NAD is a central compound in cellular metabolism. The final metabolic step in the pathway is conversion of nicotinamide adenine dinucleotide – product of NadD reaction – to NAD, a step catalyzed by the enzyme NAD synthetase (NadE). NaMN – substrate for NadD – can be formed by three different enzymatic reactions: in the *de novo* pathway from quinolinate, in Preiss-Handler salvage pathway from nicotinic acid, and in the nucleoside salvage pathway by deamination of nicotinamide mononucleotide. In bacteria, there are no known alternatives for the metabolic steps between NaMN and NAD. Mutants blocked in these steps cannot be recovered as auxotrophs since the required metabolites are not taken up by cells. In



the bacterial cells, the second substrate for NadE is ammonium, as opposed to glutamine for eukaryotes. NadE is an essential and conserved protein in the eubacterial nicotinamide adenine dinucleotide (NAD) biosynthesis pathway. Homologue of this gene identified in *Alloicoccus otitidis* is described in Example 5/ Table 4 (Seq. ID No 49). The protein encoded by the gene is set forth in Seq. ID No. 50.

#### Assays for measuring NadE function:

The *Alloicoccus otitidis* ORF- has been shown to encode, by sequence homology, nicotinamide adenine dinucleotide adenylyl synthase (NadE) (Seq. ID No. 49). NadE activity in the presence or absence of a putative inhibitory molecule of NadE activity is used to identify novel antimicrobial agents, which may be used to treat disease caused by *Alloicoccus otitidis*.

#### DISCONTINUOUS ASSAY:

In assay, NadE converts nicotinic acid adenine dinucleotide (NaAD) into nicotinamide adenine dinucleotide (NAD) in the presence of ammonia and ATP. Each PP<sub>i</sub> molecule produced by the NadE reaction can then be converted to two phosphate (P<sub>i</sub>) molecules in the presence of inorganic pyrophosphatase (PPase). The P<sub>i</sub> molecules present can then be quantitated with a malachite green reagent at 660 nm.

#### HPLC-based assay:

Enzyme activity can be measured by HPLC quantitation of the reaction products. A neutralized aliquots from the reaction described above was injected into an HPLC system utilizing a 250 x4.6 mm Supelcosil LC-18 5µm reversed-phase column. The elution conditions: 9 min at 100% buffer A (0.1 M potassium phosphate buffer, pH6.0, 6 min at up to 12% buffer B (buffer a, containing 20% methanol, 2.5 min at up to 45% buffer B, 2.5 min at up to 100% buffer B, and hold at 100% buffer B for 5.5 min. The eluate absorbance was monitored at 254 nm (1).

**Continuous assay:**

Coupled NadD-NadE assay. NadD and NadE can be detected in one continuous coupled assay. In first reaction, NadD combines nicotinic acid mononucleotide (NaMN) and adenosine triphosphate (ATP) to form nicotinic acid adenine dinucleotide (NaAD). NadE then converts NaAD into nicotinamide adenine dinucleotide (NAD) in the presence of ammonia and ATP. In the assay, the NAD product is reduced to NADH with alcohol dehydrogenase (ADH) and ethanol, thus permitting direct spectrometric detection of NADH at 340 nm wavelength. The coupled reaction above also includes inorganic pyrophosphatase (PPase) to prevent accumulation of the pyrophosphate byproduct from the consumption of ATP (this method can be use as HTS format).

NadE assay. In assay, NadE converts NaAD into nicotinamide adenine dinucleotide (NAD) in the presence of ammonia and ATP. The NAD product is reduced to NADH with alcohol dehydrogenase (ADH) and ethanol, thus permitting direct spectrometric detection of NADH at 340 nm wavelength. The reaction above also includes inorganic pyrophosphatase (PPase) to prevent accumulation of the pyrophosphate byproduct from the consumption of ATP (this method can be use as HTS format).

**EXAMPLE 35****ALLOIOCOCCUS OTITIDIS ENCODED PUTATIVE MEMBRANE PROTEIN *NOR*A**

An efflux transporter NorA that was originally identified in *Staphylococcus aureus* belongs to the family of multidrug resistance (MDR) transporters. NorA is encoded by chromosomally-located *norA* gene, it has broad substrate specificity and mediates resistance to various lipophilic and monocationic compounds such as ethidium bromide (EtBr), cetrимide, benzalkonium chloride, rhodamine 6G, tetraphenylphosphonium (TPP), chloramphenicol as well as some hygrophilic quinolones such as norfloxacin, ciprofloxacin and oxafloxacin. Increased levels of *norA* expression are associated with single nucleotide changes upstream of *norA* in a putative promoter/operator region and lead to increased pleiotropic resistance. NorA is a putative membrane protein with 12 predicted membrane-spanning domains and is classified as a member of major facilitator superfamily (MFS), a subgroup of MDR

transporters characterized by the presence of 12-14 transmembrane segments and the use of proton motive force as an energy source for drug efflux. NorA homologs that belong to MFS family include Bmr and Blt of *Bacillus subtilis*, EmeA of *Enterococcus faecalis* and PmrA of *Streptococcus pneumoniae*. The expression of *bmr* gene in *B. subtilis* is upregulated by the product of adjacent *bmr* gene in the presence of inducers (rhodamine 6G and TPP), and there is an evidence that expression of *norA* in *S. aureus* is regulated by AlrS-AlrR two-component regulatory system.

It remains unknown whether the efflux of various toxins is a primary function of NorA. When overexpressed in *E. coli*, *norA* produces resistance to a broad range of substrates including fluoroquinolones. Everted membrane vesicles prepared from *norA*-expressing *E. coli* exhibit energy-dependent transport of norfloxacin, the transfer is abolished by cyanide m-chlorophenylhydrazine (CCCP) and nigericin but not by valinomycin indicating that NorA-mediated transfer is coupled to the proton gradient of cell membrane. Norfloxacin uptake in everted vesicles as well as NorA-associated resistance phenotype is inhibited by reserpine and verapamil that also inhibit other MDR transporters and are toxic to mammalian cells. Histidine-tagged NorA (NorA-His) was recently overexpressed and purified from *E. coli*, reconstituted into both everted membrane vesicles and proteoliposomes and was shown to function as a self-sufficient efflux pump using fluorescent dye Hoechst 33342. Due to its high specificity, essentiality, and importance, *norA* is attractive as an antibacterial target. Homologue of this gene identified in *Alloiococcus otitidis* is described in Example 5/Table 4 (Seq. ID No 67). The protein encoded by the gene is set forth in Seq. ID No. 68.

#### NOR A AS A TARGET FOR ANTI-INFECTIVE DEVELOPMENT

The *Alloiococcus otitidis* ORF- has been shown to encode, by sequence homology, NorA (Seq. ID No. 67). NorA activity in the presence or absence of a putative inhibitory molecule of NorA activity is used to identify novel antimicrobial agents, which may be used to treat disease caused by *Alloiococcus otitidis*. Because of broad substrate specificity of NorA, NorA inhibitors should be particularly useful against pathogens that possess multiple drug resistance.

Whole-cell high-throughput screen (HTS) assay that measures NorA activity in the presence or absence of a putative inhibitory molecule of *Alloicoccus otitidis*. NorA activity is used to identify potential inhibitors of NorA activity. The assay utilizes *B. subtilis* strain ( $\Delta\Delta$ NA) that has both Bmr and Blt genetically inactivated while *Alloicoccus otitidis* NorA is supplied on the plasmid expression vector. The screen is based on the reversing of the resistance of  $\Delta\Delta$ NA to EtBr. The exponentially growing cells are inoculated into the wells of a 96-well plate to  $OD_{600}=0.001$ , the compounds are added at 20  $\mu$ g/ml and EtBr is added at 10  $\mu$ g/ml. Plates are incubated for 18 hrs at 37°C and examined for growth. Compounds that inhibit growth are subsequently tested in the presence/absence of EtBr for toxicity and effectivity. The efflux of EtBr from cells is monitored as described previously. The exponentially growing cells are loaded with EtBr at a concentration of 10  $\mu$ g/ml for 20 min at 37°C in the presence of reserpine (20  $\mu$ g/ml). Cells are centrifuged, resuspended to an  $OD_{600}=0.2$  in a minimal medium GM1 alone or in the presence of inhibitor compound. Fluorescence of EtBr is monitored on a fluorimeter at an excitation  $\lambda$  of 530 nm and emission  $\lambda$  of 600 nm..

#### MONITORING OF HOECHST 33342 EFFLUX

The efflux of fluorescent dye Hoechst 33342 from either everted membrane vesicles prepared from *Alloicoccus otitidis* His-NorA overexpressing *E. coli* or a proteoliposomes reconstituted with *Alloicoccus otitidis* His-NorA is also used to monitor NorA activity in the presence or absence of putative inhibitors of NorA. Everted membrane vesicles are diluted into 2 ml of 50 mM potassium HEPES (pH 7.2), 8.5 mM NaCl, 2 mM magnesium sulfate at a final protein concentration of 40  $\mu$ g/ml. NorA is activated by the addition of either 0.5 mM lactate or 0.1 mM  $Mg^{2+}$ -ATP. Hoechst 33342 is used in a range of 12.5 to 200 nM. Inhibitors are added at various concentrations prior to the addition of Hoechst 33342. Fluorescence change is monitored at excitation and emission wavelengths of 355 and 457 nm respectively in a FluoroMax spectrofluorimeter. For proteoliposome assay, the His-NorA proteoliposomes are diluted into a cuvette containing 2 ml of 20 mM potassium phosphate, 50 mM potassium sulfate, 2 mM magnesium sulfate (pH 7.0) at a protein concentration of 10  $\mu$ g/ml. The inhibitor compounds and Hoechst 33342 are added at various concentrations and the fluorescence is measured as described previously.

## EXAMPLE 36

ALLOIOCOCCUS OTITIDIS ENCODED OBG GTPASE

5       The *obg* gene is the second gene in a two-gene operon along with the stage-  
O sporulation gene *spoOB* in *B. subtilis*. SpoOB is central to the phospho-relay  
signal cascade that initiates sporulation. Obg is a member of the GTPase  
superfamily by virtue of homology throughout a small portion of the protein that in  
other members of the family is responsible for nucleotide (GTP/GDP) binding. Obg  
10 is essential for growth. Initiation of sporulation is thought to be triggered by changes  
in the GTP content of the cell; therefore, the presence of a GTP binding protein in an  
operon with a central player in the process is suggestive of a role for Obg in sensing  
GTP levels and transmitting a signal to SpoOB.

It has been shown that Obg is involved in activation of the  $\sigma^B$  transcription  
15 factor in *B. subtilis* in response to environmental stress. Cells were depleted of Obg  
utilizing a construct that put *obg* under the control of an inducible ( $P_{lac}$ ) promoter.  
Depletion of IPTG resulted in bacteria that failed to activate  $\sigma^B$ . These studies further  
showed by yeast-two-hybrid analysis that Obg interacted with several known  $\sigma^B$   
regulators, the so-called Rsb proteins.

20       The role Obg plays in transmitting signals important for sporulation and  
activation of the stress sigma factor may be indicative of the activities that small GTP  
binding proteins carry out in triggering cell division in response to GTP levels. Due to  
its high specificity, essentiality, and importance, *obg* is attractive as an antibacterial  
target. Homologue of this gene identified in *Alloiococcus otitidis* is described in  
25 Example 5/Table 4 (Seq. ID No 71). The protein encoded by the gene is set forth in  
Seq. ID No. 72.

## OBG AS A TARGET FOR ANTI-INFECTIVE DEVELOPMENT

30       Obg is essential for bacterial viability. Conditional lethal alleles revealed that  
Obg is required for early events in sporulation and is involved in transmitting signals  
require for activation of the stress sigma factor. The *Alloiococcus otitidis* ORF- has  
been shown to encode, by sequence homology, *obg* (Seq. ID No.71). Obg activity in  
the presence or absence of a putative inhibitory molecule of Obg activity is used to

identify novel antimicrobial agents, which may be used to treat disease caused by *Alloiococcus otitidis*.

#### Nucleotide binding

5 Obg binding to nucleotide in the presence or absence of putative antimicrobials, which inhibit Obg activity, is monitored by a simple filter-binding assay. *Alloiococcus otitidis* Obg (1-5  $\mu$ g) is incubated with  $\alpha^{32}$ P-GTP (0.2  $\mu$ Ci) in a buffer consisting of 50 mM Tris (pH 8.5), 1.5 mM  $\text{MgCl}_2$ , 0.1 mM EDTA, 200 mM KCl, 10% glycerol for 30 minutes to 3 hours at 37°C. A portion of the reaction mix is  
10 spotted on nitrocellulose membrane, washed (50 mM Tris (pH 8.5), 1.5 mM  $\text{MgCl}_2$ , 1 mM DTT) and dried. The membrane is then exposed to X-ray film. Alternatively, the spots are excised and counted. This assay is directly amenable to HTS using filter plates.

#### 15 GTPase activity

The GTP hydrolytic activity of Obg is monitored using thin-layer chromatography (1, 2, 10). Obg and  $\alpha^{32}$ P-GTP are incubated in 50 mM Tris (pH 8.5), 1.55 mM  $\text{MgCl}_2$ , 0.1 mM EDTA, 200 mM KCl, 10% glycerol for 30 minutes at 37°C. An aliquot of the reaction is placed on PEI cellulose and the strip developed with 0.5  
20 M  $\text{KH}_2\text{PO}_4$ , 1.0 M NaCl (pH 3.7). The spots conforming to GDP and GTP are identified by UV shadowing, excised and counted.

Alternatively, the hydrolysis of  $\gamma^{32}$ P-GTP is monitored by assaying for liberated  $\text{P}_i$  (12). Obg and  $\alpha^{32}$ P-GTP are incubated in 50 mM Tris (pH 8.5), 1.5 mM  $\text{MgCl}_2$ , 0.1 mM EDTA, 100 mM KCl, 10% glycerol for 30 minutes to 3 hours at 37°C.  
25 The reaction is stopped by the addition of a slurry of charcoal in 1 mM Kpi (pH 7.5), which selectively binds the GTP and GDP. The liberated  $\text{P}_i$  in the supernatant is monitored by Cerenkov counting. Free  $\text{P}_i$  is also monitored with the Malachite Green reagent.

#### 30 Autophosphorylation

Obg autophosphorylation is monitored by incubating Obg with  $\gamma^{32}$ P-GTP in 50 mM Tris (pH 8.5), 1.5 mM  $\text{MgCl}_2$ , 0.1 mM EDTA, 100 mM KCl, 10% glycerol for 30

minutes at 37°C. Samples are analyzed following separation on SDS polyacrylamide gels, drying the gel and exposure to film.

#### EXAMPLE 37

5     RPOA, RPOB, RPOC, AND RPOD, THE GENES ENCODING THE SUBUNITS COMPRISING  
      ALLOIOCOCCUS OTTIDIS RNA POLYMERASE: ALPHA, BETA, BETA', AND SIGMA.

RNA polymerase is an enzyme comprised of multiple highly conserved subunits which catalyzes the DNA template directed polymerization of ribonucleic nucleotides into ribonucleic acid. It is composed of a core enzyme,  $\alpha_2\beta\beta'$ , along  
10     with a fifth subunit present in stoichiometric amounts,  $\omega$  which can catalyze RNA synthesis non-specifically. Holoenzyme is formed by the introduction of the subunit  $\sigma$ , which enhances gene promoter recognition and allows specificity. Homologs of the genes identified in *Alloiococcus otitidis* are described in Example 5/Table 4 (Seq.  
15     ID Nos 7, 9, 11, and 13). The amino acid sequence of the protein encoded by these genes are set forth in Seq. ID Nos. 8, 10, 12 and 14.

Functions for the individual subunits have been defined biochemically, and interactions between them have now been deduced structurally by crystallographic analysis of the enzyme from *Thermatoga thermophila*, and to a lesser extent,  
20     *Escherichia coli*. The alpha subunit, encoded by *rpoA*, is required for enzyme assembly. It also interacts with transcription factors and with DNA elements involved in enhanced promoter strength. Beta, encoded by *rpoB*, is involved in initiation and elongation of the polymerization product. Beta' (encoded by *rpoC*), is responsible for binding of the enzyme to the DNA template. Omega is required to restore denatured  
25     RNA polymerase to function *in vitro*. Finally, sigma, encoded by *rpoD*, directs the enzyme to promoters on the template to enhance specificity of transcription (polymerization).

**ALLOIOCOCCUS OTITIDIS RNA POLYMERASE: ALPHA, BETA, BETA', AND SIGMA AS A TARGET FOR ANTI-INFECTIVE DEVELOPMENT**

Bacterial RNA polymerase is a validated target for antimicrobial chemotherapy in that several inhibitors have been identified and at least one, rifampin, is in use clinically. *Alloiococcus otitidis* RNA polymerase holoenzyme is essential for bacterial viability. The *Alloiococcus otitidis* ORFs- have been shown to encode, by sequence homology, RNA polymerase holoenzyme (Seq. ID Nos. 7, 9, 11 and 13). *Alloiococcus otitidis* RNA Polymerase activity in the presence or absence of a putative inhibitory molecule of *Alloiococcus otitidis* RNA Polymerase activity is used to identify novel antimicrobial agents, which may be used to treat disease caused by *Alloiococcus otitidis*.

**Assays for the activity of RNA polymerase**

Genes encoding the subunits of *Alloiococcus otitidis* RNA polymerase can be obtained using polymerase chain reaction amplification of the genomic region encoding them. The genes are subcloned into a standard expression vector either containing an amino acid tag for ease of purification or not. The enzyme are overexpressed in *Escherichia coli* and purified using a standard tag system or conventional chromatography .

Because RNA polymerase catalyzes the incorporation of single ribonucleotides into RNA, the incorporation of radiolabelled nucleotides into larger oligonucleotides is monitored to measure activity of the enzyme in the presence or absence of putative inhibitors of RNA polymerase activity. An automated high throughput filtration assay has been previously described for *E. coli* polymerase which uses filterplates containing a hydrophobic membrane and DEAE beads to capture polymerized RNA. G-less supercoiled DNA is used as a template at 6 ug/ml. Reaction contained 0.5 mM ATP, 0.1 mM UTP, 0.3 mM CTP, approximately 100,000 counts per minute (per 100 ul) [ $\gamma$ -<sup>33</sup>P] CTP (2000 Ci/mmol, NEN/DuPont), 4 % polyethylene glycol, 4 mM DTT, 10 mM MgCl<sub>2</sub>, in 50 mM Tris-acetate (pH 7.8), and 100 mM potassium acetate.

The reaction is carried out at 34 degrees C for 40 minutes, with 10% DMSO present in all reactions. The reaction was stopped by adding 100 ul 15% DEAE-Sephacel bead slurry in 50% methanol, 20 mM EDTA, and 0.02% NP-40. The reaction was incubated for 40-60 minutes at room temperature without shaking, and then transferred to a unifilter plate on a filtermate cell harvester. The wells were washed



six times with 2X PBS and 0.1% NP-40. After washing the bottom of the plate was sealed, and 50 ul scintillation counting liquid was added. Radioactivity was counted using a microplate scintillation counter.

Deconvolution assays are carried out by measuring the inhibition of sigma activity. Because sigma is required only for promoter specificity, polymerization may occur non-specifically if sigma is inhibited. Consequently a second assay is described above that is used to deconvolute activity against sigma.

The binding of putative inhibitory compounds to core enzyme. Several techniques are utilized to determine the interaction of inhibitors with individual subunits and include nuclear magnetic resonance and capillary electrophoresis.

#### EXAMPLE 38

##### YPHC, ENCODING A SMALL GTPASE OF UNKNOWN FUNCTION FROM *ALLOIOCOCCUS*

##### OTITIDIS

The *yphC* was initially identified in *Bacillus subtilis* in a collaboration between Wyeth and Millennium pharmaceuticals as being essential for growth by insertional mutagenesis. Subsequently it was determined that YphC, the encoded protein, contained two GTPase domains and had some homology to *era*. It was further identified in *Thermatoga maritima* and *Escherichia coli*. While no function has yet been determined for yphC, it appears that the carboxy terminal may contain an RNA binding site. In addition, site directed mutagenesis of four amino acids in the carboxy region were found to be lethal (unpublished results, Millennium). Under non-permissive conditions, strains carrying temperature sensitive alleles of the gene in *E. coli* become elongated, and chromosome segregation becomes aberrant, suggesting a role in cell division. Homologue of this gene identified in *Alloiooccus otitidis* is described in Example 5/Table 4 (Seq. ID No 73). The protein encoded by the gene is set forth in Seq. ID No. 74.

**YphC from *Alloiooccus otitidis* as a target for antimicrobial chemotherapy**

YphC is an essential protein in *Bacillus subtilis* and *E. coli*, and is conserved among bacteria including *Alloiooccus otitidis*. The *Alloiooccus otitidis* ORF- has been shown to encode, by sequence homology, YphC (Seq. ID No. 73). YphC activity in the presence or absence of a putative inhibitory molecule of YphC activity is used to identify novel antimicrobial agents, which may be used to treat disease caused by *Alloiooccus otitidis*. Consequently it is proposed here that an assay which identified inhibitors of YphC from *Alloiooccus* would result in small molecules which can be developed into effect antimicrobial agents. Additionally, because of the conservation of the enzyme among bacteria, inhibitors of the protein's function from this organism should have broad spectrum activity.

**Assays for the GTP hydrolysis by YphC**

The YphC gene from *Alloiooccus otitidis* is obtained using polymerase chain reaction amplification of the genomic region encoding it. The gene is subcloned into a standard expression vector either containing an amino acid tag for ease of purification or not. The enzyme is then overexpressed in *Escherichia coli* and purified using a standard tag system or conventional chromatography. Activity of YphC in the presence or absence putative antimicrobial agents is monitored using the assay system described below.

**GTP hydrolysis – detection by thin layer chromatography:** Reaction is carried out in a 50 ul reaction of 50 mM Tris-Cl (pH 7.5), 400 mM KCl, 5 mM MgCl<sub>2</sub>, 1 mM DTT, 10 uM [ $\alpha$ -<sup>32</sup>P] GTP, and 10 ug purified YphC, at 37 degrees for 10 minutes. The reaction is terminated by transfer of 5 ul samples to 10 ul of ice-cold 20 mM EDTA. Portions are spotted onto polyethyleneimine-cellulose thin layer chromatography plates, which are developed in 0.75 KH<sub>2</sub>PO<sub>4</sub> (pH 3.65). The plate is autoradiographed to identify hydrolysis products.

**BIBLIOGRAPHY**

- International Patent Application No. A 02322621  
International Patent Application No. EP A 125,023  
5 International Patent Application No. EP 184,187  
International Patent Application No. EP 171,496  
International Patent Application No. EP 173,494  
International Patent Application No. EP 524, 800  
International Patent Application No. WO 86/01533  
10 International Patent Application No. WO 89/02809  
International Patent Application No. WO 91/17271  
International Patent Application No. WO 92/09690  
International Patent Application No. WO 92/01047  
International Patent Application No. WO 92/18619  
15 International Patent Application No. WO 92/15679  
International Patent Application No. WO 92/20791  
International Patent Application No. WO 92/19265  
International Patent Application No. WO 93/13302  
International Patent Application No. WO 93/01288  
20 International Patent Application No. WO 97/07668  
International Patent Application No. WO 97/07669  
International Patent Application No. WO 00/18434  
PCT/US86/02269  
U.S. Patent No. 4,683195  
25 U.S. Patent No. 4,683202  
U.S. Patent No. 4,736,866  
U.S. Patent No. 4,870,009  
U.S. Patent No. 4,873,191  
U.S. Patent No. 4,987,071

- Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nuc. Acids Res.* 25(17):3389-402, 1997.
- 5 Anderson, T. F. Techniques for the preservation of three-dimensional structure in preparing specimens for the electron microscope. *Trans. N. Y. Acad. Sci.* 13(130):130-134, 1951.
- Bateman, A.T., Birney, R, Durbin, SP, Eddy, K.L. Howe, and E.L.L. Sonnhammer The Pfam protein families database *Nuc. Acids Res.* 28 (1) 263-266, 2000.
- 10 Benson, G. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res.* 27(2):573-80, 1999.
- Briles, D. E., E. Ades, J. C. Paton, J. S. Sampson, G. M. Carlone, R. C. Huebner, A. Virolainen, E. Swiatlo, and S. K. Hollingshead. Intranasal immunization of mice with a mixture of the pneumococcal proteins PsaA and PspA is highly protective against nasopharyngeal carriage of *Streptococcus pneumoniae*.
- 15 *Infect. Immun.* 68(2):796-800, 2000.
- Butler, J. C., E. D. Shapiro, and G. M. Carlone. Pneumococcal vaccines: history, current status, and future directions. *Am. J. Med.* 107(1A):69S-76S, 1999.
- Crain, M. J., W. d. Waltman, J. S. Turner, J. Yother, D. F. Talkington, L. S. McDaniel, B. M. Gray, and D. E. Brile *Alloicoccus otitidis* coccal surface protein A (PspA) is serologically highly variable and is expressed by all clinically important capsular serotypes of *Streptococcus pneumoniae*. *Infect. Immun.*
- 20 *58(10):3293-9, 1990.*
- Cserzo, M., E. Wallin, I. Simon, G. von Heijne, and A. Elofsson, Prediction of transmembrane alpha-helices in prokaryotic membrane proteins: the dense alignment surface method. *Protein Engineering* 10(6):673-6, 1997.
- 25 Dintilhac, A., G. Alloing, C. Granadel, and J.-P. Claverys. Competence and virulence of *Alloicoccus otitidis*: Adc and PsaA mutants exhibit a requirement for Zn and Mn resulting from inactivation of putative ABC metal permeases. *Mol. Microbiol.* 25(4):727-739, 1997.
- 30

- Douglas, R. M., J. C. Paton, S. J. Duncan, and D. J. Hansman. Antibody response to pneumococcal vaccination in children younger than five years of age. *J. Infect. Dis.* 148:131-137, 1983.
- 5 Eddy, S. R. Hidden Markov models. *Current Opinion in Structural Biology* 6(3):361-5, 1996.
- Foster, T. J., and M. Hook. Surface protein adhesins of *Staphylococcus aureus*. *Trends Microbiol.* 6(12):484-8, 1998.
- Fraser, C. M., S. Casjens, W. M. Huang, G. G. Sutton, R. Clayton, R. Lathigra, O.  
10 White, K. A. Ketchum, R. Dodson, E. K. Hickey, M. Gwinn, B. Dougherty, J. F. Tomb, R. D. Fleischmann, D. Richardson, J. Peterson, A. R. Kerlavage, J. Quackenbush, S. Salzberg, M. Hanson, R. van Vugt, N. Palmer, M. D. Adams, J. Gocayne, J. C. Venter, and *et al.* Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi* (see comments). *Nature*  
15 390(6660):580-6, 1997.
- Gentz *et al.*, *Proc. Natl Acad. Sci. USA*, 86, 821-824, 1989
- Goldstein, F. W., and J. Garau. 30 years of penicillin-resistant *S. pneumoniae*: myth or reality? (see comments). *Lancet* 350(9073):233-4.
- Gray, B.M., *Conjugate Vaccines Supplement* p694-697, 1990.
- 20 Hausdorff, W. P., J. Bryant, P. R. Paradiso, and G. R. Siber. 2000. Which pneumococcal serogroups cause the most invasive disease: implications for conjugate vaccine formulation and use, part I. *Clinical Infectious Diseases* 30(1):100-21, 1997.
- Hernandez-Sanchez, J., J. G. Valadez, J. V. Herrera, C. Ontiveros, and G.  
25 Guameros. lambda bar minigene-mediated inhibition of protein synthesis involves accumulation of peptidyl-tRNA and starvation for tRNA. *EMBO Jour.* 17(13):3758-65, 1998.
- Isberg, R. R., and G. Tran Van Nhieu. Binding and internalization of microorganisms by integrin receptors. *Trends in Microbiol.* 2(1):10-4, 1994.
- 30 Koebnik, R. Proposal for a peptidoglycan-associating alpha-helical motif in the C-terminal regions of some bacterial cell-surface proteins (letter; comment). *Mol. Microbiol.* 16(6):1269-70, 1995

- Lennette, D. A. General principles for laboratory diagnosis of viral, rickettsial, and chlamydial infections, p. 17-18, *diagnostic procedures for viral, rickettsial, and chlamydial infections*, vol. 7th edition, 1995.
- 5    Lewis, K. Programmed death in bacteria. *Microbiol. Mol. Biol. Rev.* 64(3):503-14, 2000.
- Loessner, M. J., S. Gaeng, and S. Scherer. Evidence for a holin-like protein gene fully embedded out of frame in the endolysin gene of Staphylococcus aureus bacteriophage 187. *J. Bacteriol.* 181(15):4452-60, 1999.
- 10    Lowry, O. H., N. J. Rosebrough, A.L. Farr and R. J. Randall. Protein measurement with the Folin-Phenol reagents. *J. Biol. Chem.* 193:265-275, 1951.
- Lukashin, A. V., and M. Borodovsky. GeneMark.hmm: new solutions for gene finding. *Nuc. Acids Res.* 26(4):1107-15, 1998.
- Mazmanian, S. K., G. Liu, H. Ton-That, and O. Schneewind.
- 15    Staphylococcus aureus sortase, an enzyme that anchors surface proteins to the cell wall. *Science* 285(5428):760-3, 1999.
- McAtee, C. P., K. E. Fry, and D. E. Berg. Identification of potential diagnostic and vaccine candidates of Helicobacter pylori by "proteome" technologies. *Helicobacter* 3(3):163-9, 1998a.
- 20    McAtee, C. P., M. Y. Lim, K. Fung, M. Velligan, K. Fry, T. Chow, and D. E. Berg. Identification of potential diagnostic and vaccine candidates of Helicobacter pylori by two-dimensional gel electrophoresis, sequence analysis, and serum profiling. *Clin. Diagn. Lab. Immunol* 5(4):537-42, 1998b.
- McAtee, C. P., M. Y. Lim, K. Fung, M. Velligan, K. Fry, T. P. Chow, and D. E. Berg.
- 25    Characterization of a Helicobacter pylori vaccine candidate by proteome techniques. *J. Chromatogr. B. Biomed. Sci. Appl.* 714(2):325-33, 1998c.
- McDaniel, L. S., G. Scott, J. F. Kearney, and D. E. Briles. Monoclonal antibodies against protease-sensitive pneumococcal antigens can protect mice from fatal infection with Streptococcus pneumoniae. *J. Exp. Med.* 160(2):386-97, 1984.
- 30    Mejlhede, N., J. F. Atkins, and J. Neuhaed. Ribosomal -1 frameshifting during decoding of Bacillus subtilis cdd occurs at the sequence CGA AAG. *J. Bacteriol.* 181(9):2930-7, 1999.

- Nabors, G. S., P. A. Braun, D. J. Hermann, M. L. Heise, D. J. Pyle, S. Gravenstein, M. Schilling, L. M. Ferguson, S. K. Hollingshead, D. E. Briles, and R. S. Becker. Immunization of healthy adults with a single recombinant pneumococcal surface protein A (PspA) variant stimulates broadly cross-reactive antibodies to heterologous PspA molecules. *Vaccine* 18:1743-1754, 2000.
- Nakai, K., and M. Kanehisa. Expert system for predicting protein localization sites in gram-negative bacteria. *Proteins* 11(2):95-110, 1991.
- 10 Navarre, W. W., and O. Schneewind. Surface Proteins of Gram-Positive Bacteria and Mechanisms of Their Targeting to the Cell Wall Envelope. *Microbiol. Mol. Biol. Rev.* 63(1):174-229, 1999.
- Nielsen, H., J. Engelbrecht, S. Brunak, and G. von Heijne. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Engineering* 10(1):1-6, 1997.
- 15 Olmsted, S. B., S. L. Erlandsen, G. M. Dunny, and C. L. Wells. High-resolution visualization by field emission scanning electron microscopy of *Enterococcus faecalis* surface proteins encoded by the pheromone-inducible conjugative plasmid pCF10. *J. Bacteriol.* 175(19):6229-37, 1993.
- 20 Park, J., and S. A. Teichmann. DIVCLUS: an automatic method in the GEANFAMMER package that finds homologous domains in single- and multi-domain proteins. *Bioinformatics* 14(2):144-50, 1998.
- Parkhill, J., M. Achtman, K. D. James, S. D. Bentley, C. Churcher, S. R. Klee, G. Morelli, D. Basham, D. Brown, T. Chillingworth, R. M. Davies, P. Davis, K. Devlin, T. Feltwell, N. Hamlin, S. Holroyd, K. Jagels, S. Leather, S. Moule, K. Mungall, M. A. Quail, M. A. Rajandream, K. M. Rutherford, M. Simmonds, J. Skelton, S. Whitehead, B. G. Spratt, and B. G. Barrell. Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491 (see comments). *Nature* 404(6777):502-6, 2000.
- 25 Pierschbacher, M. D., and E. Ruoslahti. Influence of stereochemistry of the sequence Arg-Gly-Asp-Xaa on binding specificity in cell adhesion. *J. Biol. Chem.* 262(36):17294-8, 1987.
- 30 Pizza, M., V. Scarlato, V. Masignani, M. M. Giuliani, B. Arico, M.

- Comanducci, G. T. Jennings, L. Baldi, E. Bartolini, B. Capecchi, C. L. Galeotti, E. Luzzi, R. Manetti, E. Marchetti, M. Mora, S. Nuti, G. Ratti, L. Santini, S. Savino, M. Scarselli, E. Storni, P. Zuo, M. Broeker, E. Hundt, B. Knapp, E. Blair, T. Mason, H. Tettelin, D. W. Hood, A. C. Jeffries, N. J. Saunders, D. M. Granoff, J. C. Venter, E. R. Moxon, G. Grandi, and R. Rappuoli. Identification of vaccine candidates against serogroup B meningococcus by whole-genome sequencing (see comments). *Science* 287(5459):1816-20, 2000.
- 10 Pugsley, A. P. The complete general secretory pathway in gram-negative bacteria. *Microbiol. Rev.* 57(1):50-108, 1993.
- Queen, and Baltimore, D. *Cell* 33:741-748 (1983)
- Rosenow, C., P. Ryan, J. N. Weiser, S. Johnson, P. Fontan, A. Ortqvist, and H. R. Masure. Contribution of novel choline-binding proteins to adherence, colonization and immunogenicity of *Streptococcus pneumoniae*. *Mol. Microbiol.* 25(5):819-29, 1997.
- 15 Sambrook et al., "Molecular Cloning: A Laboratory Manual: 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, NY 1989
- Salzberg, S. L., A. L. Delcher, S. Kasif, and O. White. Microbial gene identification using interpolated Markov models. *Nuc. Acids Res.* 26(2):544-8, 1998.
- 20 Sampson, J. S., S. P. O'Connor, A. R. Stinson, J. A. Tharpe, and H. Russell. Cloning and nucleotide sequence analysis of *psaA*, the *Streptococcus pneumoniae* gene encoding a 37-kilodalton protein homologous to previously reported *Streptococcus* sp. adhesins. *Infect. Immun.* 62(1):319-24, 1994.
- 25 Shinefield, H. R., and S. Black. Efficacy of pneumococcal conjugate vaccines in large scale field trials (In Process Citation). *Pediatr. Infect. Dis. J.* 19(4):394-7, 2000.
- Sonnenberg, M. G., and J. T. Belisle. Definition of *Mycobacterium tuberculosis* culture filtrate proteins by two-dimensional polyacrylamide gel electrophoresis, N-terminal amino acid sequencing, and electrospray mass spectrometry. *Infect. Immun.* 65(11):4515-24, 1997.
- 30



- Sonnhammer, E. L., G. von Heijne, and A. Krogh. A hidden Markov model for predicting transmembrane helices in protein sequences. *Ismb* 6:175-82, 1998.
- 5 Stockbauer, K. E., L. Magoun, M. Liu, E. H. Burns, Jr., S. Gubba, S. Renish, X. Pan, S. C. Bodary, E. Baker, J. Coburn, J. M. Leong, and J. M. Musser. A natural variant of the cysteine protease virulence factor of group A streptococcus with an arginine-glycine-aspartic acid (RGD) motif preferentially binds human integrins  $\alpha$ v $\beta$ 3 and  $\alpha$ 11 $\beta$ 3 (In Process Citation). *Proc. Natl. Acad. Sci. U S A* 96(1):242-7, 1999.
- 10 Talkington, D. F., B. G. Brown, J. A. Tharpe, A. Koenig, and H. Russell. Protection of mice against fatal pneumococcal challenge by immunization with pneumococcal surface adhesin A (PsaA). *Microb. Pathog.* 21(1):17-22, 1996.
- Tebbey, P. W., C. A. Unczur, J. A. Peek, D. Zhu, N. A. LaPierre, E. D. Phillips, A. R. 15 Ibraghimov, B. A. Green, J. H. Eldridge, and G. E. Hancock. Effective mucosal immunization against respiratory syncytial virus using a genetically detoxified cholera holotoxin, CT-E29H. *Vaccine* 18(24):2723-34, 2000.
- Weldingh, K., I. Rosenkrands, S. Jacobsen, P. B. Rasmussen, M. J. Elhay, and P. Andersen. 1998. Two-dimensional electrophoresis for analysis of 20 *Mycobacterium tuberculosis* culture filtrate and purification and characterization of six novel proteins. *Infect. Immun.* 66(8):3492-500, 1998.
- Yamamoto, M., D. E. Briles, S. Yamamoto, M. Ohmura, H. Kiyono, and J. R. McGhee. A nontoxic adjuvant for mucosal immunity to pneumococcal surface protein A. *J. Immunol.* 161(8):4115-21, 1998.
- 25 Calder, R. B., R. S. Williams, G. Ramaswamy, C. O. Rock, E. Campbell, S. E. Unkles, J. R. Kinghorn, and S. Jackowski. 1999. Cloning and characterization of a eukaryotic pantothenate kinase gene (panK) from *Aspergillus nidulans*. *J Biol Chem* 274:2014-20.
- Dunn, S. D., and E. E. Snell. 1979. Isolation of temperature-sensitive pantothenate 30 kinase mutants of *Salmonella typhimurium* and mapping of the *coaA* gene. *J Bacteriol* 140:805-8.

- Gerdes, S. Y., M. D. Scholle, M. D'Souza, A. Bernal, M. V. Baev, M. Farrell, O. V. Kurnasov, M. D. Daugherty, F. Mseeh, B. M. Polanuyer, J. W. Campbell, S. Anantha, K. Y. Shatalin, S. A. Chowdhury, M. Y. Fonstein, and A. L. Osterman. 2002. From genetic footprinting to antimicrobial drug targets: examples in cofactor biosynthetic pathways. *J Bacteriol* 184:4555-72.
- 5 Jackowski, S. 1996. Biosynthesis of Pantothenic acid and Coenzyme A, p. 687-694. In F. C. Neidhardt, R. Curtiss, C. A. Gross, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, R. W., M. Riley, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*. American Society for Microbiology, Washington D.C.
- 10 Rock, C. O., R. B. Calder, M. A. Karim, and S. Jackowski. 2000. Pantothenate kinase regulation of the intracellular concentration of coenzyme A. *J Biol Chem* 275:1377-83.
- 15 Song, W. J., and S. Jackowski. 1992. Cloning, sequencing, and expression of the pantothenate kinase (coaA) gene of *Escherichia coli*. *J Bacteriol* 174:6411-7.
- Song, W. J., and S. Jackowski. 1994. Kinetics and regulation of pantothenate kinase from *Escherichia coli*. *J Biol Chem* 269:27051-8.
- Vallari, D. S., S. Jackowski, and C. O. Rock. 1987. Regulation of pantothenate kinase by coenzyme A and its thioesters. *J Biol Chem* 262:2468-71.
- 20 Vallari, D. S., and C. O. Rock. 1987. Isolation and characterization of temperature-sensitive pantothenate kinase (coaA) mutants of *Escherichia coli*. *J Bacteriol* 169:5795-800.
- Espinosa-Ruiz, A., J. M. Belles, R. Serrano, and F. A. Culianez-Macia. 1999
- 25 *Arabidopsis thaliana* AtHAL3: a flavoprotein related to salt and osmotic tolerance and plant growth. *The Plant Journal: for Cell & Molecular Biology*. 20:529-539.
1. Jackowski, S. 1996. Biosynthesis of Pantothenic acid and Coenzyme A, p. 687-694. In F. C. Neidhardt, R. Curtiss, C. A. Gross, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, R. W., M. Riley, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli and Salmonella typhimurium: Cellular and*
- 30 *Molecular Biology*. American Society for Microbiology, Washington D.C.

1. Gerdes, S. Y., M. D. Scholle, M. D'Souza, A. Bernal, M. V. Baev, M. Farrell, O. V. Kurnasov, M. D. Daugherty, F. Msee, B. M. Polanuyer, J. W. Campbell, S. Anantha, K. Y. Shatalin, S. A. Chowdhury, M. Y. Fonstein, and A. L. Osterman. 2002. From genetic footprinting to antimicrobial drug targets: examples in cofactor biosynthetic pathways. *J Bacteriol* 184:4555-72.
1. Kupke, T., M. Uebele, D. Schmid, G. Jung, M. Blaesse, and S. Steinbacher. 2000. Molecular characterization of lantibiotic-synthesizing enzyme EpiD reveals a function for bacterial Dfp proteins in coenzyme A biosynthesis. *Biochemistry* 275:31838-31846.
- Kupke, T. 2001 Molecular characterization of the 4'-phosphopantothienoylcysteine decarboxylase domain of bacterial Dfp flavoproteins. *Biochemistry* 276: 27597-27604.
- Strauss, E., C. Kinsland, Y. Ge, F. W. McLafferty, and T. P. Begley. 2001. Phosphopantothienoylcysteine synthetase from *Escherichia coli*. *J. Biol. Chem.* 276:13513-13516.
- Lee, C. H., and A. F. Chen. 1982. p. 189 In J. Everse, B. Anderson, and K. You (ed), *The Pyridine Nucleotide Coenzymes*. Academic Press, Inc., New York.
- Gerdes, S. Y., M. D. Scholle, et al. 2002. From genetic footprinting to antimicrobial drug targets: examples in cofactor biosynthetic pathways. *J Bacteriol* 184: 4555-72.
- Yang, H. and R. H. Abeles. 1987. Purification and properties of *Escherichia coli* 4'-phosphopantothienoylcysteine decarboxylase: presence of covalently bound pyruvate. *Biochemistry* 26:4076-4081.
- Masuda, H., T. Uchiumi, et al. 2002. Effects of replacement of prolines with alanines on the catalytic activity and thermostability of inorganic pyrophosphatase from thermophilic bacterium PS-3. *J Biochem (Tokyo)* 131:53-8
- Lanzetta, P. A., L. J. Alvarez, P. S. Reinach, and O. A. Candia. 1979. An improved assay for nanomole amounts of inorganic phosphate. *Anal. Biochem.* 100:95-97.
- Webb, M. R. 1992. A continuous spectrophotometric assay for inorganic phosphate and for measuring phosphate release kinetics in biological systems. *PNAS* 89:4884-4887.

- Bucovaz, E. T., R. M. Macleod, J. C. Morrison, and W. D. Whybrew. 1997. The coenzyme A-synthesizing protein complex and its proposed role in CoA biosynthesis in bakers' yeast. *Biochimie* 79:787-98.
- 5 Daugherty, M., B. Polanuyer, M. Farrell, M. Scholle, A. Lykidis, V. de Crecy-Lagard, and A. Osterman. 2002. Complete reconstitution of the human coenzyme A biosynthetic pathway via comparative genomics. *J Biol Chem* 277:21431-9.
- Freiberg, C., B. Wieland, F. Spaltmann, K. Ehlert, H. Brotz, and H. Labischinski. 2001. Identification of novel essential *Escherichia coli* genes conserved among pathogenic bacteria. *J Mol Microbiol Biotechnol* 3:483-9.
- 10 Geerlof, A., A. Lewendon, and W. V. Shaw. 1999. Purification and characterization of phosphopantetheine adenylyltransferase from *Escherichia coli*. *J Biol Chem* 274:27105-11.
- Gerdes, S. Y., M. D. Scholle, M. D'Souza, A. Bernal, M. V. Baev, M. Farrell, O. V. Kurnasov, M. D. Daugherty, F. Msee, B. M. Polanuyer, J. W. Campbell, S. Anantha, K. Y. Shatalin, S. A. Chowdhury, M. Y. Fonstein, and A. L. Osterman. 2002. From genetic footprinting to antimicrobial drug targets: examples in cofactor biosynthetic pathways. *J Bacteriol* 184:4555-72.
- 15 Izard, T., and A. Geerlof. 1999. The crystal structure of a novel bacterial adenylyltransferase reveals half of sites reactivity. *Embo J* 18:2021-30.
- 20 Jackowski, S. 1996. Biosynthesis of Pantothenic acid and Coenzyme A, p. 687-694. In F. C. Neidhardt, R. Curtiss, C. A. Gross, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, R. W., M. Riley, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*. American Society for Microbiology, Washington D.C.
- 25 Lamprecht, W., and I. Trautschold. 1974. p. 2127-2131. In H. U. Bergmeyer (ed.), *Methods of Enzymatic Analysis*, 2 ed, vol. 4. Academic Press, New York.
- O'Brien, W. E. 1976. A continuous spectrophotometric assay for argininosuccinate synthetase based on pyrophosphate formation. *Anal Biochem* 76:423-30.
- 30 Worrall, D. M., and P. K. Tubbs. 1983. A bifunctional enzyme complex in coenzyme A biosynthesis: purification of pantetheine phosphate adenylyltransferase and dephospho-CoA kinase. *Biochem J* 215:153-7.

- Daugherty, M., B. Polanuyer, M. Farrell, M. Scholle, A. Lykidis, V. de Crecy-Lagard, and A. Osterman. 2002. Complete reconstitution of the human coenzyme A biosynthetic pathway via comparative genomics. *J Biol Chem* 277:21431-9.
- 5 Gerdes, S. Y., M. D. Scholle, M. D'Souza, A. Bernal, M. V. Baev, M. Farrell, O. V. Kurnasov, M. D. Daugherty, F. Msee, B. M. Polanuyer, J. W. Campbell, S. Anantha, K. Y. Shatalin, S. A. Chowdhury, M. Y. Fonstein, and A. L. Osterman. 2002. From genetic footprinting to antimicrobial drug targets: examples in cofactor biosynthetic pathways. *J Bacteriol* 184:4555-72.
- 10 Hare, R. S., S. S. Walker, T. E. Dorman, J. R. Greene, L. M. Guzman, T. J. Kenney, M. C. Sulavik, K. Baradaran, C. Houseweart, H. Yu, Z. Foldes, A. Motzer, M. Walbridge, G. H. Shimer, Jr., and K. J. Shaw. 2001. Genetic footprinting in bacteria. *J Bacteriol* 183:1694-706.
- Jackowski, S. 1996. Biosynthesis of Pantothenic acid and Coenzyme A, p. 687-694. In F. C. Neidhardt, R. Curtiss, C. A. Gross, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, R. W., M. Riley, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*. American Society for Microbiology, Washington D.C.
- 15 Jaworek, A., W. Gruber, and H. U. Bergmeyer. 1974. p. 2127-2131. In H. U. Bergmeyer (ed.), *Methods of Enzymatic Analysis*, 2 ed, vol. 4. Academic Press, New York.
- 20 Martin, D. P., and D. G. Drueckhammer. 1993. Separate enzymes catalyze the final two steps of coenzyme A biosynthesis in *Brevibacterium ammoniagenes*: purification of pantetheine phosphate adenylyltransferase. *Biochem Biophys Res Commun* 192:1155-61.
- 25 Mishra, P., P. K. Park, and D. G. Drueckhammer. 2001. Identification of *yacE* (*coaE*) as the structural gene for dephosphocoenzyme A kinase in *Escherichia coli* K-12. *J Bacteriol* 183:2774-8.
- 30 Worrall, D. M., and P. K. Tubbs. 1983. A bifunctional enzyme complex in coenzyme A biosynthesis: purification of pantetheine phosphate adenylyltransferase and dephospho-CoA kinase. *Biochem J* 215:153-7.

- Waksman, G., Lanka, E. and J. Carazo (2000) *Nature Struct. Biol.* 7:20-22.
- Gorbalenya, A.E. and E.V. Koonin (1993) *Curr. Opin. Struct. Biol.* 3:419-429.
- 5        Jezewska, M.J., Rajendran, S. and W. Bujalowski (1997) *Biochemistry*  
3:10320-10326.
- Sivaraja, M., Giordano, H. and M.G. Peterson (1998) *Anal. Biochem.* 265:22-27.
- Bruck, I. and M. O'Donnell (2000) *J. Biol. Chem.* 275:28971-28983.
- Dervyn, E., Suski, C., Daniel, R., Bruand, C., Chapuis, J., Errington, J.,
- 10        Janniere, L. and S.D. Ehrlich (2001) *Science* 294:1716-1719.
- Zechner, E.L., Wu, C.A. and K.J. Mariani (1992) *J. Biol. Chem.* 267:4054.
- Marians, K.J. (1992) *Annu. Rev. Biochem.* 61:673-719.
- Keck, J.L., Roche, D.D., Lynch, A.S. and J.M. Berger (2000) *Science*  
287:2482-2486.
- 15        Zhang, Y., Yang, F., Kao, Y., Kurilla, M.G., Pompliano, D.L., and I.B. Dicker  
(2002) *Anal. Biochem.* 304:174-179.
- Kornberg, A. and T. Baker (1992) *DNA Replication*, 2nd Ed., 169-173, W.H. Freeman  
& Co., New York
- Bruck, I. and M. O'Donnell (2000) *J. Biol. Chem.* 275:28971-28983.
- 20        O'Donnell, M. Jeruzalmi, D. and J. Kuriyan (2001) *Curr. Biol.* 11(22):R935-946.
- Yao, N., Leu, F.P., Anjelkovic, J., Turner, J. and M. O'Donnell (2000) *J. Biol. Chem.*  
275:11440-11450.
- Turner, J., Hingorani, M.M., Kelman, Z. and M. O'Donnell (1999) *EMBO* 18:771-783.
- Pellecchia, M., Sem, D.S. and K. Wuthrich (2002) *Nature Reviews. Drug Discovery.*  
25        1:211-219.
- Le, X.C., Wan, Q.H. and M. Lam (2002) *Electrophoresis* 23:903-908.
- Flurer, C.L. (2001) *Electrophoresis*. 22:4249-4261.
- Sung, W.C. and S.H. Chen (2001) *Electrophoresis* 22:4244-4248.
- Britton, R. A., S. M. Chen, D. Wallis, T. Koeuth, B. S. Powell, L. G. Shaffer, D.
- 30        Largaespada, N. A. Jenkins, N. G. Copeland, D. L. Court, and J. R. Lupski.  
2000. Isolation and preliminary characterization of the human and mouse  
homologues of the bacterial cell cycle gene *era*. *Genomics* 67:78-82.

- Britton, R. A., and J. R. Lupski. 1997. Isolation and characterization of suppressors of two *Escherichia coli* dnaG mutations, dnaG2903 and parB. *Genetics* 145:867-75.
- 5 Britton, R. A., B. S. Powell, D. L. Court, and J. R. Lupski. 1997. Characterization of mutations affecting the *Escherichia coli* essential GTPase era that suppress two temperature-sensitive dnaG alleles. *J Bacteriol* 179:4575-82.
- Britton, R. A., B. S. Powell, S. Dasgupta, Q. Sun, W. Margolin, J. R. Lupski, and D. L. Court. 1998. Cell cycle arrest in Era GTPase mutants: a potential growth rate-regulated checkpoint in *Escherichia coli*. *Mol Microbiol* 27:739-50.
- 10 Chen, S. M., H. E. Takiff, A. M. Barber, G. C. Dubois, J. C. Bardwell, and D. L. Court. 1990. Expression and characterization of RNase III and Era proteins. Products of the rnc operon of *Escherichia coli*. *J Biol Chem* 265:2888-95.
- Chen, X., D. L. Court, and X. Ji. 1999. Crystal structure of ERA: a GTPase-dependent cell cycle regulator containing an RNA binding motif. *Proc Natl Acad Sci U S A* 96:8396-401.
- 15 Inada, T., K. Kawakami, S. M. Chen, H. E. Takiff, D. L. Court, and Y. Nakamura. 1989. Temperature-sensitive lethal mutant of era, a G protein in *Escherichia coli*. *J Bacteriol* 171:5017-24.
- 20 Johnstone, B. H., A. A. Handler, D. K. Chao, V. Nguyen, M. Smith, S. Y. Ryu, E. L. Simons, P. E. Anderson, and R. W. Simons. 1999. The widely conserved Era G-protein contains an RNA-binding domain required for Era function in vivo. *Mol Microbiol* 33:1118-31.
- Lanzetta, P. A., L. J. Alvarez, P. S. Reinach, and O. A. Candia. 1979. An improved assay for nanomole amounts of inorganic phosphate. *Anal Biochem* 100:95-7.
- 25 Lerner, C. G., and M. Inouye. 1991. Pleiotropic changes resulting from depletion of Era, an essential GTP-binding protein in *Escherichia coli*. *Mol Microbiol* 5:951-7.
- March, P. E. 1992. Membrane-associated GTPases in bacteria. *Mol Microbiol* 30 6:1253-7.

- Mittenhuber, G. 2001. Comparative genomics of prokaryotic GTP-binding proteins (the Era, Obg, EngA, ThdF (TrmE), YchF and YihA families) and their relationship to eukaryotic GTP-binding proteins (the DRG, ARF, RAB, RAN, RAS and RHO families). *J Mol Microbiol Biotechnol* 3:21-35.
- 5     Sayed, A., S. Matsuyama, and M. Inouye. 1999. Era, an essential *Escherichia coli* small G-protein, binds to the 30S ribosomal subunit. *Biochem Biophys Res Commun* 264:51-4.
- 10     Sood, P., C. G. Lerner, T. Shimamoto, Q. Lu, and M. Inouye. 1994. Characterization of the autophosphorylation of Era, an essential *Escherichia coli* GTPase. *Mol Microbiol* 12:201-8.
- 15     Sullivan, S. M., R. Mishra, R. R. Neubig, and J. R. Maddock. 2000. Analysis of guanine nucleotide binding and exchange kinetics of the *Escherichia coli* GTPase Era. *J Bacteriol* 182:3460-6.
- 20     Welsh, K. M., K. A. Trach, C. Folger, and J. A. Hoch. 1994. Biochemical characterization of the essential GTP-binding protein Obg of *Bacillus subtilis*. *J Bacteriol* 176:7161-8.
- 25     Zhao, G., T. I. Meier, R. B. Peery, P. Matsushima, and P. L. Skatrud. 1999. Biochemical and molecular analyses of the C-terminal domain of Era GTPase from *Streptococcus pneumoniae*. *Microbiology* 145 ( Pt 4):791-800.
- 30     Bouhss, A., Josseaume, N., Allan, D., Crouvoisier, M., Gutmann, L., Mainardi, J.-L., Mengin-Lecreulx, D., van Heijenoort, J., and M. Arthur. 2001. Identification of the UDP-MurNAc-pentapeptide:L-alanine ligase for synthesis of branched peptidoglycan precursors in *Enterococcus faecalis*. *J. Bacteriol.* 183: 5122-5127.
- Berger-Bachi, B., Barberis-Maino, L., Strassle, A., and F. H. Kayser. 1989. FemA, a host-mediated factor essential for methicillin resistance in *Staphylococcus aureus*: molecular cloning and characterization. *Mol. Gen. Genet.* 219: 263-269.
- Flouret, B., Mengin-Lecreulx, D., and J. van Heijenoort. 1981. Reverse-phase high-pressure liquid chromatography of uridine diphosphate N-acetylmuramyl peptide precursors of bacterial cell wall peptidoglycan. *Anal. Biochem.* 114: 59-63.



- Filipe, S. R., Pinho, M. G., and A. Tomasz. 2000. Characterization of the murMN operon involved in the synthesis of branched peptidoglycan peptides in *Streptococcus pneumoniae*. *J. Biol. Chem.* 275: 27768-27774.
- 5 Filipe, S. R., and A. Tomasz. 2000. Inhibition of the expression of penicillin resistance in *Streptococcus pneumoniae* by inactivation of cell wall muropeptide branching genes. *Proc. Natl. Acad. Sci. USA* 97: 4891-4896
- Hegde, S. S., and T. E. Shrader. 2001. FemABX family members are novel nonribosomal peptidyltransferases and important pathogen-specific drug
- 10 targets. *J. Biol. Chem.* 276:6998-7003.
- de Jonge, B. L., Sidow, T., Chang, Y. S., Labischinski, H., Berger-Bachi, B., Gage, D. A., and A. Tomasz. 1993. Altered muropeptide composition in *Staphylococcus aureus* strains with an inactivated femA locus. *J. Bacteriol.* 175: 2779-2782.
- 15 Kopp, U., Roos, M., Wecke, J., and H. Labischinski. 1996. Staphylococcal peptidoglycan interpeptide bridge biosynthesis: a novel antistaphylococcal target? *Microb. Drug Resist.* 2: 29-41.
- de Lencastre, H., and A. Tomasz. 1994. Reassessment of the number of auxiliary genes essential for expression of high-level methicillin resistance in
- 20 *Staphylococcus aureus*. *Antimicrob. Agents Chemother.* 38: 2590-2598.
- Matsushashi, M., Dietrich, C. P., and J. L. Strominger. 1967. Biosynthesis of the peptidoglycan of bacterial cell wall. III. The role of soluble ribonucleic acid and of lipid intermediates in glycine incorporation in *Staphylococcus aureus*. *J. Biol. Chem.* 242: 3191-3206.
- 25 Rohrer, S., Ehlert, K., Tschierske, M., Labischinski, H., and B. Berger-Bachi. 1999. The essential *Staphylococcus aureus* gene fmhB is involved in the first step of peptidoglycan pentaglycine interpeptide formation. *Proc. Natl. Acad. Sci. USA* 96: 9351-9356.
- Ruzin, A., Severin, A., Ritacco, F., Tabei, K., Singh, G., Bradford, P. A., Siegel, M.
- 30 M., Projan, S. J., and D. M. Shlaes. 2002 Further evidence that a cell wall precursor [C<sub>55</sub>-MurNAc-(peptide)-GlcNAc] serves as an acceptor in a sorting reaction. *J. Bacteriol.* 184: 2141-2147.

- Stranden, A. M., Ehler, K., Labischinski, H., and B. Berger-Bachi. 1997. Cell wall monoglycine cross-bridges and methicillin hypersusceptibility in a femAB null mutant of methicillin-resistant *Staphylococcus aureus*. *J. Bacteriol.* 179: 9-16.
- 5 Dale, G.E., R.L. Then, and D. Stuber, Characterization of the gene for chromosomal trimethoprim-sensitive dihydrofolate reductase of *Staphylococcus aureus* ATCC 25923. *Antimicrob Agents Chemother*, 1993. 37(7): p. 1400-5.
- Brown, G.M. and J.M. Williamson, Biosynthesis of folic acid, riboflavin, thiamine and pantothenic acid, in *Escherichia coli* and *Salmonella typhimurium*: Cellular and Molecular Biology, F.C. Neidhardt, et al., Editors. 1987, American Society for Microbiology: Washington, D.C. p. 521-538.
- 10 Xiao, B., et al., Crystal structure of 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase, a potential target for the development of novel antimicrobial agents. *Structure Fold Des*, 1999. 7(5): p. 489-96.
- 15 Hennig, M., et al., Crystal structure and reaction mechanism of 7,8-dihydroneopterin aldolase from *Staphylococcus aureus*. *Nat Struct Biol*, 1998. 5(5): p. 357-62.
- Schweitzer, B.I., A.P. Dicker, and J.R. Bertino, Dihydrofolate reductase as a therapeutic target. *Faseb J*, 1990. 4(8): p. 2441-52.
- Li, R., et al., Three-dimensional structure of *M. tuberculosis* dihydrofolate reductase reveals opportunities for the design of novel tuberculosis drugs. *J Mol Biol*, 2000. 295(2): p. 307-23.
- 20 Quinlivan, E.P., et al., Mechanism of the antimicrobial drug trimethoprim revisited. *FASEB Journal*, 2000. 14(15): p. 2519-24.
- Li, L.Y. and S.J. Benkovic, Impact on catalysis of secondary structural manipulation of the alpha C-helix of *Escherichia coli* dihydrofolate reductase. *Biochemistry*, 1991. 30(6): p. 1470-8.
- 25 Hennig, M., et al., Crystal structure and reaction mechanism of 7,8-dihydroneopterin aldolase from *Staphylococcus aureus*. *Nat Struct Biol*, 1998. 5(5): p. 357-62.
- Brown, G.M. and J.M. Williamson, Biosynthesis of folic acid, riboflavin, thiamine and pantothenic acid, in *Escherichia coli* and *Salmonella typhimurium*: Cellular and Molecular Biology, F.C. Neidhardt, et al., Editors. 1987, American Society for Microbiology: Washington, D.C. p. 521-538.
- 30

- Mouillon, J.M., et al., Folate synthesis in higher-plant mitochondria: coupling between the dihydropterin pyrophosphokinase and the dihydropteroate synthase activities. *Biochem J*, 2002. 363(Pt 2): p. 313-9.
- 5 Lopez, P., et al., Sulfonamide resistance in *Streptococcus pneumoniae*: DNA sequence of the gene encoding dihydropteroate synthase and characterization of the enzyme. *J Bacteriol*, 1987. 169(9): p. 4320-6.
- Lopez, P., B. Greenberg, and S.A. Lacks, DNA sequence of folate biosynthesis gene *sulD*, encoding hydroxymethyldihydropterin pyrophosphokinase in
- 10 *Streptococcus pneumoniae*, and characterization of the enzyme. *J Bacteriol*, 1990. 172(9): p. 4766-74.
- Margolis, P.S., A. Driks, and R. Losick, Sporulation gene *spoIIB* from *Bacillus subtilis*. *J Bacteriol*, 1993. 175(2): p. 528-40.
- Brown, G.M. and J.M. Williamson, Biosynthesis of folic acid, riboflavin, thiamine and
- 15 pantothenic acid, in *Escherichia coli* and *Salmonella typhimurium*: Cellular and Molecular Biology, F.C. Neidhardt, et al., Editors. 1987, American Society for Microbiology: Washington, D.C. p. 521-538.
- Hennig, M., et al., Crystal structure and reaction mechanism of 7,8-dihydroneopterin aldolase from *Staphylococcus aureus*. *Nat Struct Biol*, 1998. 5(5): p. 357-62.
- 20 Xiao, B., et al., Crystal structure of 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase, a potential target for the development of novel antimicrobial agents. *Structure Fold Des*, 1999. 7(5): p. 489-96.
- Mouillon, J.M., et al., Folate synthesis in higher-plant mitochondria: coupling between the dihydropterin pyrophosphokinase and the dihydropteroate synthase
- 25 activities. *Biochem J*, 2002. 363(Pt 2): p. 313-9.
- Hampele, I.C., et al., Structure and function of the dihydropteroate synthase from *Staphylococcus aureus*. *J Mol Biol*, 1997. 268(1): p. 21-30.
- Bognar, A.L. and B. Shane, Bacterial folylpoly(gamma-glutamate) synthase-dihydrofolate synthase. *Methods Enzymol*, 1986. 122: p. 349-59.
- 30 Ferone, R. and A. Warskow, Co-purification of dihydrofolate synthetase and N10formyltetrahydropteroyldiglutamate synthetase from *E. coli*. *Adv Exp Med Biol*, 1983. 163: p. 167-81.

- Fussenegger, M. and T.F. Meyer, Cloning and characterization of the *Neisseria gonorrhoeae* MS11 folC gene. *Mol Gen Genet*, 1996. 250(3): p. 277-85.
- Pyne, C. and A.L. Bognar, Replacement of the folC gene, encoding  
5 folypolyglutamate synthetase-dihydrofolate synthetase in *Escherichia coli*,  
with genes mutagenized in vitro. *J Bacteriol*, 1992. 174(6): p. 1750-9.
- Garrow, T.A., A. Admon, and B. Shane, Expression cloning of a human cDNA  
encoding folypoly(gamma-glutamate) synthetase and determination of its  
primary structure. *Proc Natl Acad Sci U S A*, 1992. 89(19): p. 9151-5.
- 10 Webb, M.R., A continuous spectrophotometric assay for inorganic phosphate and for  
measuring phosphate release kinetics in biological systems. *Proc Natl Acad  
Sci U S A*, 1992. 89(11): p. 4884-7.
- Lanzetta, P.A., et al., An improved assay for nanomole amounts of inorganic  
phosphate. *Anal Biochem*, 1979. 100(1): p. 95-7.
- 15 Welsh, K.M., et al., Biochemical characterization of the essential GTP-binding protein.  
Og of *Bacillus subtilis*. *J Bacteriol*, 1994. 176(23): p. 7161-8.
- Margolis, P.S., A. Driks, and R. Losick, Sporulation gene *spoIIB* from *Bacillus subtilis*.  
*J Bacteriol*, 1993. 175(2): p. 528-40.
- Brown, G.M. and J.M. Williamson, Biosynthesis of folic acid, riboflavin, thiamine and  
20 pantothenic acid, in *Escherichia coli* and *Salmonella typhimurium*: Cellular  
and Molecular Biology, F.C. Neidhardt, et al., Editors. 1987, American Society  
for Microbiology: Washington, D.C. p. 521-538.
- Hennig, M., et al., Crystal structure and reaction mechanism of 7,8-dihydroneopterin  
aldolase from *Staphylococcus aureus*. *Nat Struct Biol*, 1998. 5(5): p. 357-62.
- 25 Xiao, B., et al., Crystal structure of 6-hydroxymethyl-7,8-dihydropterin  
pyrophosphokinase, a potential target for the development of novel  
antimicrobial agents. *Structure Fold Des*, 1999. 7(5): p. 489-96.
- Mouillon, J.M., et al., Folate synthesis in higher-plant mitochondria: coupling between  
the dihydropterin pyrophosphokinase and the dihydropteroate synthase  
30 activities. *Biochem J*, 2002. 363(Pt 2): p. 313-9.
- Hampele, I.C., et al., Structure and function of the dihydropteroate synthase from  
*Staphylococcus aureus*. *J Mol Biol*, 1997. 268(1): p. 21-30.

- Bognar, A.L. and B. Shane, Bacterial folypoly(gamma-glutamate) synthase-dihydrofolate synthase. *Methods Enzymol*, 1986. 122: p. 349-59.
- 5 Ferone, R. and A. Warskow, Co-purification of dihydrofolate synthetase and N10formyltetrahydropteroyldiglutamate synthetase from *E. coli*. *Adv Exp Med Biol*, 1983. 163: p. 167-81.
- Fussenegger, M. and T.F. Meyer, Cloning and characterization of the *Neisseria gonorrhoeae* MS11 folC gene. *Mol Gen Genet*, 1996. 250(3): p. 277-85.
- 10 Pyne, C. and A.L. Bognar, Replacement of the folC gene, encoding folypolyglutamate synthetase-dihydrofolate synthetase in *Escherichia coli*, with genes mutagenized in vitro. *J Bacteriol*, 1992. 174(6): p. 1750-9.
- Garrow, T.A., A. Admon, and B. Shane, Expression cloning of a human cDNA encoding folypoly(gamma-glutamate) synthetase and determination of its primary structure. *Proc Natl Acad Sci U S A*, 1992. 89(19): p. 9151-5.
- 15 Webb, M.R., A continuous spectrophotometric assay for inorganic phosphate and for measuring phosphate release kinetics in biological systems. *Proc Natl Acad Sci U S A*, 1992. 89(11): p. 4884-7.
- Lanzetta, P.A., et al., An improved assay for nanomole amounts of inorganic phosphate. *Anal Biochem*, 1979. 100(1): p. 95-7.
- 20 Welsh, K.M., et al., Biochemical characterization of the essential GTP-binding protein Ogb of *Bacillus subtilis*. *J Bacteriol*, 1994. 176(23): p. 7161-8.
- Xiao, B., et al., Crystal structure of 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase, a potential target for the development of novel antimicrobial agents. *Structure Fold Des*, 1999. 7(5): p. 489-96.
- 25 Hennig, M., et al., The structure and function of the 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase from *Haemophilus influenzae*. *J Mol Biol*, 1999. 287(2): p. 211-9.
- Brown, G.M. and J.M. Williamson, Biosynthesis of folic acid, riboflavin, thiamine and pantothenic acid, in *Escherichia coli* and *Salmonella typhimurium*: Cellular and Molecular Biology, F.C. Neidhardt, et al., Editors. 1987, American Society for Microbiology: Washington, D.C. p. 521-538.
- 30

- Akerley, B.J., et al., A genome-scale analysis for identification of genes required for growth or survival of *Haemophilus influenzae*. *Proc Natl Acad Sci U S A*, 2002. 99(2): p. 966-71.
- 5 Mouillon, J.M., et al., Folate synthesis in higher-plant mitochondria: coupling between the dihydropterin pyrophosphokinase and the dihydropteroate synthase activities. *Biochem J*, 2002. 363(Pt 2): p. 313-9.
- Illarionova, V., et al., Biosynthesis of Tetrahydrofolate. *STEREOCHEMISTRY OF DIHYDRONEOPTERIN ALDOLASE*. *J Biol Chem*, 2002. 277(32): p. 28841-7.
- 10 Lopez, P., et al., Sulfonamide resistance in *Streptococcus pneumoniae*: DNA sequence of the gene encoding dihydropteroate synthase and characterization of the enzyme. *J Bacteriol*, 1987. 169(9): p. 4320-6.
- Lopez, P., B. Greenberg, and S.A. Lacks, DNA sequence of folate biosynthesis gene *sulD*, encoding hydroxymethyldihydropterin pyrophosphokinase in
- 15 *Streptococcus pneumoniae*, and characterization of the enzyme. *J Bacteriol*, 1990. 172(9): p. 4766-74.
- Slock, J., et al., An apparent *Bacillus subtilis* folic acid biosynthetic operon containing *pab*, an amphibolic *trpG* gene, a third gene required for synthesis of para-aminobenzoic acid, and the dihydropteroate synthase gene. *J Bacteriol*, 1990.
- 20 172(12): p. 7211-26.
- Hampele, I.C., et al., Structure and function of the dihydropteroate synthase from *Staphylococcus aureus*. *J Mol Biol*, 1997. 268(1): p. 21-30.
- Brown, G.M. and J.M. Williamson, Biosynthesis of folic acid, riboflavin, thiamine and pantothenic acid, in *Escherichia coli* and *Salmonella typhimurium*: Cellular and Molecular Biology, F.C. Neidhardt, et al., Editors. 1987, American Society
- 25 for Microbiology: Washington, D.C. p. 521-538.
- Hennig, M., et al., Crystal structure and reaction mechanism of 7,8-dihydroneopterin aldolase from *Staphylococcus aureus*. *Nat Struct Biol*, 1998. 5(5): p. 357-62.
- Crystal structure of the anti-bacterial sulfonamide drug target dihydropteroate
- 30 synthase. *Nat Struct Biol*, 1997. 4(6): p. 490-7.
- Akerley, B.J., et al., A genome-scale analysis for identification of genes required for growth or survival of *Haemophilus influenzae*. *Proc Natl Acad Sci U S A*, 2002. 99(2): p. 966-71.

- Mouillon, J.M., et al., Folate synthesis in higher-plant mitochondria: coupling between the dihydropterin pyrophosphokinase and the dihydropteroate synthase activities. *Biochem J*, 2002. 363(Pt 2): p. 313-9.
- 5 Vinnicombe, H.G. and J.P. Derrick, Dihydropteroate synthase from *Streptococcus pneumoniae*: characterization of substrate binding order and sulfonamide inhibition. *Biochem Biophys Res Commun*, 1999. 258(3): p. 752-7.
- Rebeille, F., et al., Folate biosynthesis in higher plants: purification and molecular cloning of a bifunctional 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase/7,8-dihydropteroate synthase localized in mitochondria.
- 10 *Embo J*, 1997. 16(5): p. 947-57.
- Pucci, M.J., et al., Identification and characterization of cell wall-cell division gene clusters in pathogenic gram-positive cocci. *J Bacteriol*, 1997. 179(17): p. 5632-5.
- 15 Lutkenhaus, J. and S.G. Addinall, Bacterial cell division and the Z ring. *Annu Rev Biochem*, 1997. 66: p. 93-116.
- Rothfield, L., S. Justice, and J. Garcia-Lara, Bacterial cell division. *Annu Rev Genet*, 1999. 33: p. 423-48.
- Ricard, M. and Y. Hirota, Process of cellular division in *Escherichia coli*: physiological study on thermosensitive mutants defective in cell division. *J Bacteriol*, 1973.
- 20 116(1): p. 314-22.
- Lutkenhaus, J.F., H. Wolf-Watz, and W.D. Donachie, Organization of genes in the *ftsA-envA* region of the *Escherichia coli* genetic map and identification of a new *fts* locus (*ftsZ*). *J Bacteriol*, 1980. 142(2): p. 615-20.
- 25 Robinson, A.C., et al., DNA sequence and transcriptional organization of essential cell division genes *ftsQ* and *ftsA* of *Escherichia coli*: evidence for overlapping transcriptional units. *J Bacteriol*, 1984. 160(2): p. 546-55.
- Hale, C.A. and P.A. de Boer, Direct binding of *FtsZ* to *ZipA*, an essential component of the septal ring structure that mediates cell division in *E. coli*. *Cell*, 1997.
- 30 88(2): p. 175-85.
- Hale, C.A. and P.A. de Boer, Recruitment of *ZipA* to the septal ring of *Escherichia coli* is dependent on *FtsZ* and independent of *FtsA*. *J Bacteriol*, 1999. 181(1): p. 167-76.

- Wang, X., et al., Analysis of the interaction of FtsZ with itself, GTP, and FtsA. *J Bacteriol*, 1997. 179(17): p. 5551-9.
- van den Ent, F. and J. Lowe, Crystal structure of the cell division protein FtsA from *Thermotoga maritima*. *Embo J*, 2000. 19(20): p. 5300-7.
- 5 Pichoff, S. and J. Lutkenhaus, Unique and overlapping roles for ZipA and FtsA in septal ring assembly in *Escherichia coli*. *Embo J*, 2002. 21(4): p. 685-93.
- Akerley, B.J., et al., A genome-scale analysis for identification of genes required for growth or survival of *Haemophilus influenzae*. *Proc Natl Acad Sci U S A*, 10 2002. 99(2): p. 966-71.
- Beall, B. and J. Lutkenhaus, Impaired cell division and sporulation of a *Bacillus subtilis* strain with the *ftsA* gene deleted. *J Bacteriol*, 1992. 174(7): p. 2398-403.
- Dewar, S.J., K.J. Begg, and W.D. Donachie, Inhibition of cell division initiation by an imbalance in the ratio of FtsA to FtsZ. *J Bacteriol*, 1992. 174(19): p. 6314-6.
- 15 Dai, K. and J. Lutkenhaus, The proper ratio of FtsZ to FtsA is required for cell division to occur in *Escherichia coli*. *J Bacteriol*, 1992. 174(19): p. 6145-51.
- Bork, P., C. Sander, and A. Valencia, An ATPase domain common to prokaryotic cell cycle proteins, sugar kinases, actin, and hsp70 heat shock proteins. *Proc Natl Acad Sci U S A*, 1992. 89(16): p. 7290-4.
- 20 Feucht, A., et al., Cytological and biochemical characterization of the FtsA cell division protein of *Bacillus subtilis*. *Mol Microbiol*, 2001. 40(1): p. 115-25.
- Ma, X., et al., Interactions between heterologous FtsA and FtsZ proteins at the FtsZ ring. *J Bacteriol*, 1997. 179(21): p. 6788-97.
- 25 Yan, K., K.H. Pearce, and D.J. Payne, A conserved residue at the extreme C-terminus of FtsZ is critical for the FtsA-FtsZ interaction in *Staphylococcus aureus*. *Biochem Biophys Res Commun*, 2000. 270(2): p. 387-92.
- Najafi, S.M., D.A. Harris, and M.D. Yudkin, The SpoIIAA protein of *Bacillus subtilis* has GTP-binding properties. *J Bacteriol*, 1996. 178(22): p. 6632-4.



- Haney, S.A., et al., Genetic analysis of the *Escherichia coli* FtsZ.ZipA interaction in the yeast two-hybrid system. Characterization of FtsZ residues essential for the interactions with ZipA and with FtsA. *J Biol Chem*, 2001. 276(15): p. 11980-7.
- 5 Addinall, S. G., E. Bi, and J. Lutkenhaus. 1996. FtsZ ring formation in fts mutants. *J Bacteriol* 178:3877-84.
- Addinall, S. G., C. Cao, and J. Lutkenhaus. 1997. FtsN, a late recruit to the septum in *Escherichia coli*. *Mol Microbiol* 25:303-9.
- 10 Addinall, S. G., and J. Lutkenhaus. 1996. FtsA is localized to the septum in an FtsZ-dependent manner. *J Bacteriol* 178:7167-72.
- Bernander, R. 2000. Chromosome replication, nucleoid segregation and cell division in archaea. *Trends Microbiol* 8:278-83.
- 15 Bi, E. F., and J. Lutkenhaus. 1991. FtsZ ring structure associated with division in *Escherichia coli*. *Nature* 354:161-4.
- Bork, P., C. Sander, and A. Valencia. 1992. An ATPase domain common to prokaryotic cell cycle proteins, sugar kinases, actin, and hsp70. heat shock proteins. *Proc Natl Acad Sci U S A* 89:7290-4.
- 20 de Boer, P., R. Crossley, and L. Rothfield. 1992. The essential bacterial cell-division protein FtsZ is a GTPase. *Nature* 359:254-6.
- Din, N., E. M. Quardokus, M. J. Sackett, and Y. V. Brun. 1998. Dominant C-terminal deletions of FtsZ that affect its ability to localize in *Caulobacter* and its interaction with FtsA. *Mol Microbiol* 27:1051-63.
- Erickson, H. P. 1995. FtsZ, a prokaryotic homolog of tubulin? *Cell* 80:367-70.
- 25 Erickson, H. P., D. W. Taylor, K. A. Taylor, and D. Bramhill. 1996. Bacterial cell division protein FtsZ assembles into protofilament sheets and minirings, structural homologs of tubulin polymers. *Proc Natl Acad Sci U S A* 93:519-23.
- 30 Feucht, A., I. Lucet, M. D. Yudkin, and J. Errington. 2001. Cytological and biochemical characterization of the FtsA cell division protein of *Bacillus subtilis*. *Mol Microbiol* 40:115-25.
- Hale, C. A., and P. A. de Boer. 1997. Direct binding of FtsZ to ZipA, an essential component of the septal ring structure that mediates cell division in *E. coli*. *Cell* 88:175-85.

- Hale, C. A., and P. A. de Boer. 1999. Recruitment of ZipA to the septal ring of *Escherichia coli* is dependent on FtsZ and independent of FtsA. *J Bacteriol* 181:167-76.
- 5 Hale, C. A., A. C. Rhee, and P. A. de Boer. 2000. ZipA-induced bundling of FtsZ polymers mediated by an interaction between C-terminal domains. *J Bacteriol* 182:5153-66.
- Haney, S. A., E. Glasfeld, C. Hale, D. Keeney, Z. He, and P. de Boer. 2001. Genetic analysis of the *Escherichia coli* FtsZ.ZipA interaction in the yeast two-hybrid system. Characterization of FtsZ residues essential for the interactions with ZipA and with FtsA. *J Biol Chem* 276:11980-7.
- 10 Hoenig, M., R. J. Lee, and D. C. Ferguson. 1989. A microtiter plate assay for inorganic phosphate. *J Biochem Biophys Methods* 19:249-51.
- Justice, S. S., J. Garcia-Lara, and L. I. Rothfield. 2000. Cell division inhibitors SulA and MinC/MinD block septum formation at different steps in the assembly of the *Escherichia coli* division machinery. *Mol Microbiol* 37:410-23.
- 15 Lanzetta, P. A., L. J. Alvarez, P. S. Reinach, and O. A. Candia. 1979. An improved assay for nanomole amounts of inorganic phosphate. *Anal Biochem* 100:95-7.
- Lowe, J., and L. A. Amos. 1998. Crystal structure of the bacterial cell-division protein FtsZ. *Nature* 391:203-6.
- 20 Lowe, J., and L. A. Amos. 1999. Tubulin-like protofilaments in Ca<sup>2+</sup>-induced FtsZ sheets. *EMBO J* 18:2364-71.
- Lu, C., M. Reedy, and H. P. Erickson. 2000. Straight and curved conformations of FtsZ are regulated by GTP hydrolysis. *J Bacteriol* 182:164-70.
- 25 Ma, X., D. W. Ehrhardt, and W. Margolin. 1996. Colocalization of cell division proteins FtsZ and FtsA to cytoskeletal structures in living *Escherichia coli* cells by using green fluorescent protein. *Proc Natl Acad Sci U S A* 93:12998-3003.
- Margolin, W. 2000. Themes and variations in prokaryotic cell division. *FEMS Microbiol Rev* 24:531-48.
- 30 Mosyak, L., Y. Zhang, E. Glasfeld, S. Haney, M. Stahl, J. Seehra, and W. S. Somers. 2000. The bacterial cell-division protein ZipA and its interaction with an FtsZ fragment revealed by X-ray crystallography. *EMBO J* 19:3179-91.

- Moy, F. J., E. Glasfeld, L. Mosyak, and R. Powers. 2000. Solution structure of ZipA, a crucial component of *Escherichia coli* cell division. *Biochemistry*. 39:9146-56.
- Mukherjee, A., K. Dai, and J. Lutkenhaus. 1993. *Escherichia coli* cell division protein FtsZ is a guanine nucleotide binding protein. *Proc Natl Acad Sci U S A* 90:1053-7.
- Mukherjee, A., and J. Lutkenhaus. 1999. Analysis of FtsZ assembly by light scattering and determination of the role of divalent metal cations. *J Bacteriol* 181:823-32.
- 10 Mukherjee, A., and J. Lutkenhaus. 1998. Dynamic assembly of FtsZ regulated by GTP hydrolysis. *EMBO J* 17:462-9.
- Mukherjee, A., and J. Lutkenhaus. 1994. Guanine nucleotide-dependent assembly of FtsZ into filaments. *J Bacteriol* 176:2754-8.
- Pichoff, S., and J. Lutkenhaus. 2002. Unique and overlapping roles for ZipA and FtsA in septal ring assembly in *Escherichia coli*. *EMBO J* 21:685-93.
- 15 Pla, J., M. Sanchez, P. Palacios, M. Vicente, and M. Aldea. 1991. Preferential cytoplasmic location of FtsZ, a protein essential for *Escherichia coli* septation. *Mol Microbiol* 5:1681-6.
- RayChaudhuri, D., and J. T. Park. 1992. *Escherichia coli* cell-division gene *ftsZ* encodes a novel GTP-binding protein. *Nature* 359:251-4.
- 20 Romberg, L., M. Simon, and H. P. Erickson. 2001. Polymerization of FtsZ, a bacterial homolog of tubulin. is assembly cooperative? *J Biol Chem* 276:11743-53.
- Rothfield, L., S. Justice, and J. Garcia-Lara. 1999. Bacterial cell division. *Annu Rev Genet* 33:423-48.
- 25 Sanchez, M., A. Valencia, M. J. Ferrandiz, C. Sander, and M. Vicente. 1994. Correlation between the structure and biochemical activities of FtsA, an essential cell division protein of the actin family. *EMBO J* 13:4919-25.
- Scheffers, D. J., J. G. de Wit, T. den Blaauwen, and A. J. Driessen. 2002. GTP hydrolysis of cell division protein FtsZ: evidence that the active site is formed by the association of monomers. *Biochemistry (Mosc)* 41:521-9.
- 30 Scheffers, D. J., J. G. de Wit, T. den Blaauwen, and A. J. Driessen. 2001. Substitution of a conserved aspartate allows cation-induced polymerization of FtsZ. *FEBS Lett* 494:34-7.

- Scheffers, D. J., T. den Blaauwen, and A. J. Driessen. 2000. Non-hydrolysable GTP-  
gamma-S stabilizes the FtsZ polymer in a GDP-bound state. *Mol Microbiol*  
35:1211-9.
- 5 Stricker, J., P. Maddox, E. D. Salmon, and H. P. Erickson. 2002. Rapid assembly  
dynamics of the Escherichia coli FtsZ-ring demonstrated by fluorescence  
recovery after photobleaching. *Proc Natl Acad Sci U S A* 99:3171-5.
- Wang, X., J. Huang, A. Mukherjee, C. Cao, and J. Lutkenhaus. 1997. Analysis of the  
interaction of FtsZ with itself, GTP, and FtsA. *J Bacteriol* 179:5551-9.
- 10 Yan, K., K. H. Pearce, and D. J. Payne. 2000. A conserved residue at the extreme C-  
terminus of FtsZ is critical for the FtsA-FtsZ interaction in Staphylococcus  
aureus. *Biochem Biophys Res Commun* 270:387-92.
- Ali, J. A., A. P. Jackson, A. J. Howells, and A. Maxwell. 1993. The 43-kilodalton N-  
terminal fragment of the DNA gyrase B protein hydrolyzes ATP and binds  
15 coumarin drugs. *Biochemistry* 32:2717-24.
- Baker, T. A., B. E. Funnell, and A. Kornberg. 1987. Helicase action of dnaB protein  
during replication from the Escherichia coli chromosomal origin in vitro. *J Biol*  
*Chem* 262:6877-85.
- Barnabe, N., and B. B. Hasinoff. 2001. High-throughput fluorescence flow-injection  
20 topoisomerase II inhibition assay. *J Chromatogr B Biomed Sci Appl* 760:263-  
9.
- Blanche, F., B. Cameron, F. X. Bernard, L. Maton, B. Manse, L. Ferrero, N. Ratet, C.  
Lecoq, A. Goniot, D. Bisch, and J. Crouzet. 1996. Differential behaviors of  
Staphylococcus aureus and Escherichia coli type II DNA topoisomerases.  
25 *Antimicrob Agents Chemother* 40:2714-20.
- Boehm, H. J., M. Boehringer, D. Bur, H. Gmuender, W. Huber, W. Klaus, D.  
Kostrewa, H. Kuehne, T. Luebbers, N. Meunier-Keller, and F. Mueller. 2000.  
Novel inhibitors of DNA gyrase: 3D structure based biased needle screening,  
hit validation by biophysical methods, and 3D guided optimization. A  
30 promising alternative to random screening. *J Med Chem* 43:2664-74.
- Caron, P. R., and J. C. Wang. 1994. Apperidix. II: Alignment of primary sequences of  
DNA topoisomerases. *Adv Pharmacol* 29B:271-97.

- Chatterji, M., S. Unniraman, A. Maxwell, and V. Nagaraja. 2000. The additional 165 amino acids in the B protein of *Escherichia coli* DNA gyrase have an important role in DNA binding. *J Biol Chem* 275:22888-94.
- 5 Ferrero, L., B. Cameron, and J. Crouzet. 1995. Analysis of *gyrA* and *grlA* mutations in stepwise-selected ciprofloxacin-resistant mutants of *Staphylococcus aureus*. *Antimicrob Agents Chemother* 39:1554-8.
- Ferrero, L., B. Cameron, B. Manse, D. Lagneaux, J. Crouzet, A. Famechon, and F. Blanche. 1994. Cloning and primary structure of *Staphylococcus aureus* DNA  
10 topoisomerase IV: a primary target of fluoroquinolones. *Mol Microbiol* 13:641-53.
- Forsyth, R. A., R. J. Haselbeck, K. L. Ohlsen, R. T. Yamamoto, H. Xu, J. D. Trawick, D. Wall, L. Wang, V. Brown-Driver, J. M. Froelich, K. G. C, P. King, M. McCarthy, C. Malone, B. Misiner, D. Robbins, Z. Tan, Z. Y. Zhu Zy, G. Carr,  
15 D. A. Mosca, C. Zamudio, J. G. Foulkes, and J. W. Zyskind. 2002. A genome-wide strategy for the identification of essential genes in *Staphylococcus aureus*. *Mol Microbiol* 43:1387-400.
- Fuller, R. S., J. M. Kaguni, and A. Kornberg. 1981. Enzymatic replication of the origin of the *Escherichia coli* chromosome. *Proc Natl Acad Sci U S A* 78:7370-4.
- 20 Gmunder, H., K. Kuratli, and W. Keck. 1995. Effect of pyrimido[1,6-a]benzimidazoles, quinolones, and Ca<sup>2+</sup> on the DNA gyrase-mediated cleavage reaction. *Antimicrob Agents Chemother* 39:163-9.
- Kaguni, J. M., and A. Kornberg. 1984. Replication initiated at the origin (*oriC*) of the *E. coli* chromosome reconstituted with purified enzymes. *Cell* 38:183-90.
- 25 Lassota, P., G. Singh, and R. Kramer. 1996. Mechanism of topoisomerase II inhibition by staurosporine and other protein kinase inhibitors. *J Biol Chem* 271:26418-23.
- Lewis, R. J., O. M. Singh, C. V. Smith, T. Skarzynski, A. Maxwell, A. J. Wonacott, and D. B. Wigley. 1996. The nature of inhibition of DNA gyrase by the  
30 coumarins and the cyclothialidines revealed by X-ray crystallography. *Embo J* 15:1412-20.

- Lockshon, D., and D. R. Morris. 1985. Sites of reaction of *Escherichia coli* DNA gyrase on pBR322 in vivo as revealed by oxolinic acid-induced plasmid linearization. *J Mol Biol* 181:63-74.
- 5 Manjunatha, U. H., M. Dalal, M. Chatterji, D. R. Radha, S. S. Visweswariah, and V. Nagaraja. 2002. Functional characterisation of mycobacterial DNA gyrase: an efficient decatenase. *Nucleic Acids Res* 30:2144-53.
- Maxwell, A. 1997. DNA gyrase as a drug target. *Trends Microbiol* 5:102-9.
- Maxwell, A. 1993. The interaction between coumarin drugs and DNA gyrase. *Mol*  
10 *Microbiol* 9:681-6.
- Menzel, R., and M. Gellert. 1994. The biochemistry and biology of DNA gyrase. *Adv Pharmacol* 29A:39-69.
- Morais Cabral, J. H., A. P. Jackson, C. V. Smith, N. Shikotra, A. Maxwell, and R. C. Liddington. 1997. Crystal structure of the breakage-reunion domain of DNA  
15 gyrase. *Nature* 388:903-6.
- Nakada, N., H. Gmunder, T. Hirata, and M. Arisawa. 1995. Characterization of the binding site for cyclothialidine on the B subunit of DNA gyrase. *J Biol Chem* 270:14286-91.
- Nakada, N., H. Shimada, T. Hirata, Y. Aoki, T. Kamiyama, J. Watanabe, and M.  
20 Arisawa. 1993. Biological characterization of cyclothialidine, a new DNA gyrase inhibitor. *Antimicrob Agents Chemother* 37:2656-61.
- Pellecchia, M., D. S. Sem, and K. Wuthrich. 2002. NMR in drug discovery. *Nat Rev Drug Discov* 1:211-9.
- Peng, H., and K. J. Marians. 1993. Decatenation activity of topoisomerase IV during  
25 *oriC* and pBR322 DNA replication in vitro. *Proc Natl Acad Sci U S A* 90:8571-5.
- Peng, H., and K. J. Marians. 1993. *Escherichia coli* topoisomerase IV. Purification, characterization, subunit structure, and subunit interactions. *J Biol Chem* 268:24481-90.
- 30 Reece, R. J., and A. Maxwell. 1991. The C-terminal domain of the *Escherichia coli* DNA gyrase A subunit is a DNA-binding protein. *Nucleic Acids Res* 19:1399-405.

- Reece, R. J., and A. Maxwell. 1991. DNA gyrase: structure and function. *Crit Rev Biochem Mol Biol* 26:335-75.
- Reece, R. J., and A. Maxwell. 1989. Tryptic fragments of the Escherichia coli DNA  
5 gyrase A protein. *J Biol Chem* 264:19648-53.
- Sahai, B. M., and J. G. Kaplan. 1986. A quantitative decatenation assay for type II  
topoisomerases. *Anal Biochem* 156:364-79.
- Sung, W. C., and S. H. Chen. 2001. Recent advances in pharmacokinetic  
applications of capillary electrophoresis. *Electrophoresis* 22:4244-8.
- 10 Wang, J. C. 1991. DNA topoisomerases: why so many? *J Biol Chem* 266:6659-62.
- Wigley, D. B. 1995. Structure and mechanism of DNA topoisomerases. *Annu Rev  
Biophys Biomol Struct* 24:185-208.
- Wigley, D. B., G. J. Davies, E. J. Dodson, A. Maxwell, and G. Dodson. 1991. Crystal  
structure of an N-terminal fragment of the DNA gyrase B protein. *Nature*  
15 351:624-9.
- Willmott, C. J., S. E. Critchlow, I. C. Eperon, and A. Maxwell. 1994. The complex of  
DNA gyrase and quinolone drugs with DNA forms a barrier to transcription by  
RNA polymerase. *J Mol Biol* 242:351-63.
- 7Zechiedrich, E. L., and N. R. Cozzarelli. 1995. Roles of topoisomerase IV and DNA  
20 gyrase in DNA unlinking during replication in Escherichia coli. *Genes Dev*  
9:2859-69.
- Brown ED. Marquardt JL. Lee JP. Walsh CT. Anderson KS. 1994. Detection and  
characterization of a phospholactoyl-enzyme adduct in the reaction catalyzed  
by UDP-N-acetylglucosamine enolpyruvoyl transferase, MurZ. *Biochemistry*,  
25 33: 10638-45.
- Brown E. Zhao X. 1997. Biochemical inhibitor capture screen for inhibitors of UDP-  
N-GluNac enolpyruvyltransferase (MurA/Z). Millennium Pharmaceuticals Inc.  
Standard Operating Procedure.
- Brown DE. Vives EL. Walsh CT. Kolter R. 1995. MurA (MurZ) the enzyme that  
30 catalyzes the first committed step in peptidoglycan biosynthesis, is essential  
in Escherichia coli. *J. Bacteriol.* 177: 4194-97.

- Bugg TDH. Walsh CT. 1992. Interacellular steps of bacteria cell wall peptidoglycan biosynthesis: enzymology, antibiotics and antibiotic resistance. Nat. Prod. Rep. 9: 199-215.
- 5    Dai HJ. Parker CN and Bao JJ. 2002. Characterization and inhibition study of MurA enzyme by capillary electrophoresis. J. Chromatography B: Analytical Technologies in the Biomedical & Life Sciences. 766:123-32.
- Du W. Brown JR. Sylvester DR. Huang J. Chalker AF. So CY. Holmes DJ. Payne DJ and Wallis NG. 2000. Two active forms of UDP-N-acetylglucosamine enolpyruvyl transferase in gram-positive bacteria. J Bacteriol. 182:4146-52.
- 10    Holtje JV and Schwarz U. 1985. In fosfomycin (phosphonomycin). Ann. N.Y. Academic science 235:364-85.
- Kim DH. Lees WJ. Kempell KE. Lane WS. Duncan K and Walsh CT. 1996. Characterization of a Cys115 to Asp substitution in the Escherichia coli cell wall biosynthetic enzyme UDP-GlcNAc enolpyruvyl transferase (MurA) that
- 15    confers resistance to inactivation by the antibiotic fosfomycin. Biochemistry. 35:4923-8.
- Marquardt JL. Siegele DA. Kolter R and Walsh CT 1992. Cloning and sequencing of Escherichia coli murZ and purification of its product, a UDP-N-acetylglucosamine enolpyruvyl transferase J. Bacteriol. 174: 5748-52.
- 20    Nanninga (ed) Molecular cytology of Escherichia coli, p. 77-109. Academic press, New York, NY.
- Kahan FM. Kahan JS. Cassidy PJ and Krapp H (1974). The mechanism of action of Marquardt JL. Brown ED. Lane WS. Haley TM. Ichikawa Y. Wong C-H and Walsh CT. 1994. Kinetics, stoichiometry, and identification of the reactive thiolate in the
- 25    inactivation of UDP-GlcNAc enolpyruvyl transferase by the antibiotic fosfomycin. Biochemistry, 33:10646-51.
- Schonbrunn E. Svergun DI. Amrhein N and Koch MH. 1998. Studies on the conformational changes in the bacterial cell wall biosynthetic enzyme UDP-N-acetylglucosamine enolpyruvyltransferase (MurA). European J Biochem.
- 30    253(2):406-12.



- Skarzynski T. Mysstry A. Wanacott A. Hutchinson SE. Kelly VA and Duncan K. 1996.  
Structure of UDP-N-acetylglucosamine enolpyruvyl transferase, an enzyme  
essential for the synthesis of bacterial peptidoglycan, complexed with  
substrate UDP-N-acetylglucosamine and the drug fosfomycin. *Structure*  
4:1465-74.
- Samland AK. Jelesarov I. Kuhn R. Amrhein N. Macheroux P. 2001. Thermodynamic  
characterization of ligand-induced conformational changes in UDP-N-  
acetylglucosamine enolpyruvyl transferase. *Biochemistry*. 40(33):9950-6.
- Singh G and Yang Y. 1999. Wyeth-Ayerst Research, Procedure 3407.
- Andres CJ. Bronson JJ. D'Andrea SV. Deshpande MS. Falk PJ. Grant-Young KA.  
Harte WE. Ho HT. Misco PF. Robertson JG. Stock D. Sun Y. Walsh AW.  
2000. 4-Thiazolidinones: novel inhibitors of the bacterial enzyme MurB.  
*Bioorganic & Medicinal Chemistry Letters*. 10: 715-7.
- Benson TE. Walsh CT. Hogle JM. The structure of the substrate-free form of MurB,  
an essential enzyme for the synthesis of bacterial cell walls. 1996. *Structure*.  
4(1):47-54.
- Benson TE. Walsh CT. Hogle JM. X-ray crystal structures of the S229A mutant and  
wild-type MurB in the presence of the substrate enolpyruvyl-UDP-N-  
acetylglucosamine at 1.8-A resolution. 1997. *Biochemistry*. 36(4):806-11.
- Benson TE. Walsh CT. Massey V. Kinetic characterization of wild-type and S229A  
mutant MurB: evidence for the role of Ser 229 as a general acid. 1997.  
*Biochemistry*. 36(4):796-805.
- Benson TE. Harris MS. Choi GH. Cialdella JI. Herberg JT. Martin JP Jr. Baldwin ET.  
A structural variation for MurB: X-ray crystal structure of *Staphylococcus*  
*aureus* UDP-N-acetylenolpyruvylglucosamine reductase (MurB). 2000.  
*Biochemistry*. 40 :2340-50..
- Dhalla A. M., Yanchunas J., Ho H-T., Falk P. J., Villafranca J. J. and Robertson J. G.  
(1995). Steady-state kinetic mechanism of *Escherichia coli* UDP-N-  
Acetylenolpyruvylglucosamine reductase. *Biochemistry*, 34: 5390-5402.

- Constantine KL. Mueller L. Goldfarb V. Wittekind M. Metzler WJ. Yanchunas J Jr.  
Robertson JG. Malley MF. Friedrichs MS. Farmer BT 2nd. Characterization of  
NADP<sup>+</sup> binding to perdeuterated MurB: backbone atom NMR assignments  
and chemical-shift changes. 1997. *Journal of Molecular Biology*. 267(5):1223-  
46.
- DeCenzo M. T. and Henderson B. 1997. MurB Biochemical Assay. Millennium  
Pharmaceuticals Inc. Standard Procedure.
- Farmer BT 2nd. Constantine KL. Goldfarb V. Friedrichs MS. Wittekind M. Yanchunas  
J Jr. Robertson JG. Mueller L. 1996. Localizing the NADP<sup>+</sup> binding site on the  
MurB enzyme by NMR. *Nature Structural Biology*. 3:995-7.
- Krishnamurthy G. 1999. Wyeth-Ayerst Research, Notebook L-20868.
- Sarver RW. Rogers JM. Epps DE. 2002. Determination of ligand-MurB interactions  
by isothermal denaturation: application as a secondary assay to complement  
high throughput screening. *Journal of Biomolecular Screening*. 7:21-8.
- Sylvester DR. Alvarez E. Patel A. Ratnam K. Kallender H. Wallis NG. 2001.  
Identification and characterization of UDP-N-acetylenolpyruvylglucosamine  
reductase (MurB) from the Gram-positive pathogen *Streptococcus*  
*pneumoniae*. *Biochemical Journal*. 355(Pt 2):431-5.
- Yang Y. 1999. MurB Biochemistry Assay. Wyeth-Ayerst Research Procedure 3408.  
Yang Y. 2002. MurB (UDP-N-acetylenolpyruvylglucosamine reductase) Inhibitor  
Screen Using a Coupled Biochemical Reaction with MurC Wyeth Research,  
Procedure 7075.
- Anderson MS. Eveland SS. Onishi HR. Pompliano DL. 1996. Kinetic mechanism of  
the *Escherichia coli* UDPMurNAc-tripeptide D-alanyl-D-alanine-adding  
enzyme: use of a glutathione S-transferase fusion. *Biochemistry*.  
35(50):16264-9.
- Bouhss A. Mengin-Lecreulx D. Blanot D. van Heijenoort J. Parquet C. 1997.  
Invariant amino acids in the Mur peptide synthetases of bacterial  
peptidoglycan synthesis and their modification by site-directed mutagenesis in  
the UDP-MurNAc:L-alanine ligase from *Escherichia coli*. *Biochemistry*.  
36:11556-63.

- El Zoeiby A. Sanschagrin F. Lamoureux J. Darveau A. Levesque RC. 2000. Cloning, over-expression and purification of *Pseudomonas aeruginosa* murC encoding uridine diphosphate N-acetylmuramate: L-alanine ligase. FEMS Microbiology Letters. 183(2):281-8.
- 5 Emanuele JJ Jr. Jin H. Jacobson BL. Chang CY. Einspahr HM. Villafranca JJ. 1996. Kinetic and crystallographic studies of *Escherichia coli* UDP-N-acetylmuramate:L-alanine ligase. Protein Science. 5(12):2566-74.
- Jin H. Emanuele JJ Jr. Fairman R. Robertson JG. Hail ME. Ho HT. Falk PJ. 10 Villafranca JJ. 1996. Structural studies of *Escherichia coli* UDP-N-acetylmuramate:L-alanine ligase. Biochemistry. 35(5):1423-31.
- Lanzetta PA. Alvarez JL. Reinach PS. Candia OA. 1979. An improved assay for nanomole amounts of inorganic phosphate. Analytical Biochemistry 100: 95-97.
- 15 Liger D. Masson A. Blanot D. van Heijenoort J. Parquet C. 1996. Study of the overproduced uridine-diphosphate-N-acetylmuramate:L-alanine ligase from *Escherichia coli*. Microbial Drug Resistance-Mechanisms Epidemiology & Disease. 2(1):25-7.
- Lowe AM. Deresiewicz RL. Cloning and sequencing of *Staphylococcus aureus* murC, 20 a gene essential for cell wall biosynthesis. DNA Sequence. 10(1):19-23, 1999
- Marmor S. Petersen CP. Reck F. Yang W. Gao N. Fisher SL. 2001. Biochemical characterization of a phosphinate inhibitor of *Escherichia coli* MurC. Biochemistry. 40(40):12207-14.
- Reck F. Marmor S. Fisher S. Wuonola MA. 2001. Inhibitors of the bacterial cell wall 25 biosynthesis enzyme MurC. Bioorganic & Medicinal Chemistry Letters. 11(11):1451-4.
- Schleifer KH. Kandler O. 1972. Peptidoglycan types of bacteria cell wall and their taxonomic implications. Bacteriol. Rev. 36: 407-477.
- Severin A. 1999 MurC biochemistry assay Wyeth research Procedure Number 3427.
- 30 Bertrand JA. Auger G. Fanchon E. Martin L. Blanot D. van Heijenoort J. Dideberg O. 1997. Crystal structure of UDP-N-acetylmuramoyl-L-alanine:D-glutamate ligase from *Escherichia coli*. EMBO Journal. 16(12):3416-25.

- Bertrand JA. Auger G. Martin L. Fanchon E. Blanot D. Le Beller D. van Heijenoort J. Dideberg O. 1999. Determination of the MurD mechanism through crystallographic analysis of enzyme complexes. *Journal of Molecular Biology*. 289(3):579-90.
- 5        Bertrand JA. Fanchon E. Martin L. Chantalat L. Auger G. Blanot D. van Heijenoort J. Dideberg O. 2000. "Open" structures of MurD: domain movements and structural similarities with folypolyglutamate synthetase. *Journal of Molecular Biology*. 301:1257-66.
- 10      Bouhss A. Dementin S. Parquet C. Mengin-Lecreulx D. Bertrand JA. Le Beller D. Dideberg O. van Heijenoort J. Blanot D. 1999. Role of the ortholog and paralog amino acid invariants in the active site of the UDP-MurNAc-L-alanine:D-glutamate ligase (MurD). *Biochemistry*. 38:12240-7.
- 15      El-Sherbeini M. Geissler WM. Pittman J. Yuan X. Wong KK. Pompliano DL. 1998. Cloning and expression of *Staphylococcus aureus* and *Treptococcus pyogenes* murD genes encoding uridine diphosphate N-acetylmuramoyl-L-alanine:D-glutamate ligases. *Gene*. 210:117-25.
- 20      Gegnas LD. Waddell ST. Chabin RM. Reddy S. Wong KK. 1998. Inhibitors of the bacterial cell wall biosynthesis enzyme MurD. *Bioorganic & Medicinal Chemistry Letters*. 8:1643-8.
- 25      Gobec S. Urleb U. Auger G. Blanot D. 2001. Synthesis and biochemical evaluation of some novel N-acyl phosphono- and phosphinoalanine derivatives as potential inhibitors of the D-glutamic acid-adding enzyme. *Pharmazie*. 56:295-7.
- 30      Gordon E. Flouret B. Chantalat L. van Heijenoort J. Mengin-Lecreulx D. Dideberg O. 2001. Crystal structure of UDP-N-acetylmuramoyl-L-alanyl-D-glutamate: meso-diaminopimelate ligase from *Escherichia coli*. *J. Biological Chemistry*. 276:10999-1006.
- 30      Lanzetta PA. Alvarez JL. Reinach PS. Candia OA. 1979. An improved assay for nanomole amounts of inorganic phosphate. *Analytical Biochemistry* 100: 95-97.

- Mengin-Lecreux D. van Heijenoort J. 1990. Nucleotide sequence of the murD gene encoding the UDP-MurNAc-L-Ala-D-Glu synthetase of Escherichia coli. Nucleic Acids Research. 18(1):183.
- 5   Pratviel-Sosa F. Mengin-Lecreux D. van Heijenoort J. 1991. Over-production, purification and properties of the uridine diphosphate N-acetylmuramoyl-L-alanine:D-glutamate ligase from Escherichia coli. European Journal of Biochemistry. 202:1169-76.
- 10   Vaganay S. Tanner ME. van Heijenoort J. Blanot D. 1996. Study of the reaction mechanism of the D-glutamic acid-adding enzyme from Escherichia coli. Microbial Drug Resistance. 2:51-4.
- Walsh AW. Falk PJ. Thanassi J. Discotto L. Pucci MJ. Ho HT. Comparison of the D-glutamate-adding enzymes from selected gram-positive and gram-negative bacteria. Journal of Bacteriology. 181(17):5395-401, 1999
- 15   Schleifer KH. Kandler O. 1972. Peptidoglycan types of bacteria cell wall and their taxonomic implications. Bacteriol. Rev. 36: 407-477.
- Severin A. 1999 MurD biochemistry assay Wyeth research Procedure Number 3429.
- Anderson JS. Matsushashi M. Haskin MA. Strominger JL. 1967. Biosynthesis of the peptidoglycan of bacterial cell walls. II. Phospholipid carriers in the reaction
- 20   sequence. J. Biological Chemistry 242: 3180-90.
- Bouhss A. Mengin-Lecreux D. Blanot D. van Heijenoort J. Parquet C. 1997. Invariant amino acids in the Mur peptide synthetases of bacterial peptidoglycan synthesis and their modification by site-directed mutagenesis in the UDP-MurNAc:L-alanine ligase from Escherichia coli. Biochemistry. 36(39):11556-
- 25   63.
- Chakraborty A. Das S. Majumdar S. Mukhopadhyay K. Roychoudhury S. Chaudhuri K. 2000. Use of RNA arbitrarily primed-PCR fingerprinting to identify Vibrio cholerae genes differentially expressed in the host following infection. Infection & Immunity. 68(7):3878-87.

- Dementin S. Bouhss A. Auger G. Parquet C. Mengin-Lecreulx D. Dideberg O. van Heijenoort J. Blanot D. 2001. Evidence of a functional requirement for a carbamoylated lysine residue in MurD, MurE and MurF synthetases as established by chemical rescue experiments. *European Journal of Biochemistry*. 268(22):5800-7.
- Dietrich CP. Colucci AV. Strominger JK. 1967. Biosynthesis of the peptidoglycan of bacterial cell walls. V. Separation of protein and lipid components of the particulate enzyme from *Micrococcus lysodieticus* and purification of the endogenous lipid acceptors. *J. Biological Chemistry* 242: 3218-25.
- Jana M. Luong TT. Komatsuzawa H. Shigeta M. Lee CY. 2000. A method for demonstrating gene essentiality in *Staphylococcus aureus*. *Plasmid*. 44:100-4.
- Liger D. Masson A. Blanot D. Van Heijenoort J. Parquet C. 1995. Over-production, purification and properties of the uridine-diphosphate-N-acetylmuramate:L-alanine ligase from *Escherichia coli*. *European J. Biochemistry*. 230: 80-87.
- Ludovice AM. Wu SW. de Lencastre H. 1998. Molecular cloning and DNA sequencing of the *Staphylococcus aureus* UDP-N-acetylmuramyl tripeptide synthetase (murE) gene, essential for the optimal expression of methicillin resistance. *Microbial Drug Resistance*. 4:85-90.
- Ma C. Redfield RJ. 2000. Point mutations in a peptidoglycan biosynthesis gene cause competence induction in *Haemophilus influenzae*. *Journal of Bacteriology*. 182(12):3323-30.
- Mengin-Lecreulx D. Falla T. Blanot D. van Heijenoort J. Adams DJ. Chopra I. 1999. Expression of the *Staphylococcus aureus* UDP-N-acetylmuramoyl-L-alanyl-D-glutamate:L-lysine ligase in *Escherichia coli* and effects on peptidoglycan biosynthesis and cell growth. *Journal of Bacteriology*. 181(19):5909-14.
- Michaus C. Mengin-Lecreulx D. Van Heijenoort J. Blanot D. 1990. Overproduction, purification and properties of the UDP-N-acetylmuramoyl-L-alanyl-D-glutamate: meso-diaminopimelic acid ligase from *Escherichia coli*. *Euro J. Biochem*. 194:853-61.

- Severin A. 1999 MurE biochemistry assay Wyeth research Procedure Number 3430.
- Wijayarathna CD. Wachi M. Nagai K. 2001. Isolation of *ftsI* and *murE* genes involved  
in peptidoglycan synthesis from *Corynebacterium glutamicum*. *Applied*  
5 *Microbiology & Biotechnology*. 55(4):466-70
- Anderson MS. Eveland SS. Onishi HR. Pompliano DL. 1996. Kinetic mechanism of  
the *Escherichia coli* UDP-MurNAc-tripeptide D-alanyl-D-alanine-adding  
enzyme: use of a glutathione S-transferase fusion. *Biochemistry*. 35:16264-9.
- Duncan K. van Heijenoort J. Walsh CT. 1990. Purification and characterization of the  
10 D-alanyl-D-alanine-adding enzyme from *Escherichia coli*. *Biochemistry*. 29(9):  
2379-86.
- Dementin S. Bouhss A. Auger G. Parquet C. Mengin-Lecreux D. Dideberg O. van  
Heijenoort J. Blanot D. 2001. Evidence of a functional requirement for a  
carbamoylated lysine residue in MurD, MurE and MurF synthetases as  
15 established by chemical rescue experiments. *European Journal of*  
*Biochemistry*. 268:5800-7.
- Malakhov MP. Los DA. Wada H. Semenenko VE. Murata N. 1995. Characterization  
of the *murF* gene of the cyanobacterium *Synechocystis* sp. PCC 6803.  
*Microbiology*. 141 (Pt 1): 163-9.
- 20 Nicola, GW. 1999. TITLE MurF JOURNAL Patent: JP 1999253176-A 4 21-SEP-  
1999; SmithKline BEECHAM CORP, SmithKline BEECHAM CORP PUBLIC  
LTD CO COMMENT OS *Staphylococcus aureus*.
- Nicola, GW. TITLE MurF JOURNAL Patent: JP 1999225780-A 4 24-AUG-1999;  
SmithKline BEECHAM CORP, SMITHKLINE BEECHAM CORP PUBLIC LTD  
25 CO COMMENT OS *Streptococcus pneumonia*
- Parquet C. Flouter B. Mengin-Lecreux D. van Heijenoort J. 1998. Nucleotide  
sequence of the *murF* gene encoding the UDP-MurNAc-pentapeptide  
synthetase of *Escherichia coli*. *Nucleic Acids Research*. 17(13): 5379.
- Yan Y. Munshi S. Li Y. Pryor KA. Marsilio F. Leiting B. 1999. Crystallization and  
30 preliminary X-ray analysis of the *Escherichia coli* UDP-MurNAc-tripeptide D-  
alanyl-D-alanine-adding enzyme (MurF). *Acta Crystallographica Section D-  
Biological Crystallography*. 55 ( Pt 12):2033-4

- Yan Y. Munshi S. Leiting B. Anderson MS. Chrzas J. Chen Z. 2000. Crystal structure of *Escherichia coli* UDPMurNAc-tripeptide d-alanyl-d-alanine-adding enzyme (MurF) at 2.3 Å resolution. *J Mol. Biol.* 304(3): 435-45.
- 5     Anderson JS. Matsubashi M. Haskin MA. Strominger JL. 1967. Biosynthesis of the peptidoglycan of bacterial cell walls. II. Phospholipid carriers in the reaction sequence. *J. Biol. Chem.* 242: 3180-3190.
- Baden LR. Critchley IA. Sahm DF. So W. Gedde M. Porter S. Moellering RC Jr. Eliopoulos G. 2002. Molecular characterization of vancomycin-resistant
- 10     Enterococci repopulating the gastrointestinal tract following treatment with a novel glycolipodepsipeptide, ramoplanin. *J Clin. Microb.* 40(4):1160-3.
- Branstrom AA. Midha S. Longley CB. Han K. Baizman ER. Axelrod HR. 2000. Assay for identification of inhibitors for bacterial MraY translocase or MurG transferase. *Analytical Biochemistry.* 280(2):315-9.
- 15     Chen L. Men H. Ha S. Ye XY. Brunner L. Hu Y. Walker S. 2002. Intrinsic lipid preferences and kinetic mechanism of *Escherichia coli* MurG. *Biochemistry.* 41(21):6824-33.
- Cudic P. Behenna DC. Yu MK. Kruger RG. Szewczuk LM. McCafferty DG. 2001. Synthesis of P(1)-Citronellyl-P(2)-α-D-pyranosyl pyrophosphates as
- 20     potential substrates for the *E. coli* undecaprenyl-pyrophosphoryl-N-acetylglucosaminyl transferase MurG. *Bioorganic & Medicinal Chemistry Letters.* 11(24):3107-10.
- Crouvoisier M. Mengin-Lecreux D. van Heijenoort J. 1999. UDP-N-acetylglucosamine:N-acetylmuramoyl-(pentapeptide) pyrophosphoryl
- 25     undecaprenol N-acetylglucosamine transferase from *Escherichia coli*: overproduction, solubilization, and purification. *FEBS Letters.* 449(2-3):289-92.
- Ha S. Walker D. Shi Y. Walker S. 2000. The 1.9 Å crystal structure of *Escherichia coli* MurG, a membrane-associated glycosyltransferase involved in
- 30     peptidoglycan biosynthesis. *Protein Science.* 9(6):1045-52.
- Landman D. Quale JM. Burney S. Kreiswirth B. Willey BM. 1996. Treatment of experimental endocarditis caused by multidrug resistant *Enterococcus faecium* with ramoplanin and penicillin. *J. Antimicrob. Chemother.* 37(2):323-9.



- Men H. Park P. Ge M. and Walker S. 1998. Substrate synthesis and activity assay for MurG. *J. Am. Chem. Soc.* 120:2484-85.
- 5 Mengin-Lecreulx D. Texier L. Rousseau, M. and van Heijenoort J. 1991. The murG gene of *Escherichia coli* codes for the UDP-N-acetylglucosamine:N-acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol N-acetylglucosamine transferase involved in the membrane steps of peptidoglycan synthesis. *J. Bacteriol.* 173:4625-36.
- 10 Mulichak AM. Losey HC. Walsh CT. Garavito RM. 2001. Structure of the UDP-glucosyltransferase GtfB that modifies the heptapeptide aglycone in the biosynthesis of vancomycin group antibiotics. *Structure.* 9(7):547-57.
- Rolston KV. Dholakia N. Ho DH. LeBlanc B. Dvorak T. Streeter H. 1996. In-vitro activity of ramoplanin (a novel lipoglycopeptide), vancomycin, and teicoplanin against gram-positive clinical isolates from cancer patients. *J Antimicrob. Chemother.* 38(2):265-9.
- 15 Salmond GP. Lutkenhaus, JF and Donachie W. D. 1980. Identification of new genes in a cell envelop-cell division gene cluster of *Escherichia coli*: cell envelope gene MurG. *J. Bacteriology*, 144:438-440.
- Silva DJ. Bowe CL. Branstrom AA. Baizman ER. Sofia MJ. 2000. Synthesis and biological evaluation of analogues of bacterial lipid I. *Bioorganic & Medicinal Chemistry Letters.* 10(24):2811-3.
- 20 Beach, M. J., and V. W. Rodwell. 1989. Cloning, sequencing, and overexpression of mvaA, which encodes *Pseudomonas mevalonii* 3-hydroxy-3-methylglutaryl coenzyme A reductase. *J Bacteriol* 171:2994-3001.
- 25 Goldstein, J. L., and M. S. Brown. 1990. Regulation of the mevalonate pathway. *Nature* 343:425-30.
- Hedi, M., A. Sutherlin, E. I. Wilding, M. Mazzulla, D. McDevitt, P. Lane, J. W. Burgner, 2nd, K. R. Lehnbeuter, C. V. Stauffacher, M. N. Gwynn, and V. W. Rodwell. 2002. *Enterococcus faecalis* acetoacetyl-coenzyme A thiolase/3-hydroxy-3-methylglutaryl-coenzyme A reductase, a dual-function protein of isopentenyl diphosphate biosynthesis. *J Bacteriol* 184:2116-22.
- 30 Horbach, S., H. Sahm, and R. Welle. 1993. Isoprenoid biosynthesis in bacteria: two different pathways? *FEMS Microbiol Lett* 111:135-40.

- Istvan, E. S. 2001. Bacterial and mammalian HMG-CoA reductases: related enzymes with distinct architectures. *Curr Opin Struct Biol* 11:746-51.
- Istvan, E. S., M. Palnitkar, S. K. Buchanan, and J. Deisenhofer. 2000. Crystal  
5 structure of the catalytic portion of human HMG-CoA reductase: insights into regulation of activity and catalysis. *Embo J* 19:819-30.
- Johnson, E. A., and W. A. Schroeder. 1996. Microbial carotenoids. *Adv Biochem Eng Biotechnol* 53:119-78.
- Lange, B. M., T. Rujan, W. Martin, and R. Croteau. 2000. Isoprenoid biosynthesis:  
10 the evolution of two ancient and distinct pathways across genomes. *Proc Natl Acad Sci U S A* 97:13172-7.
- Meganathan, R. 1996. Biosynthesis of the isoprenoid quinones menaquinone (vitamin K<sub>2</sub>) and ubiquinone (coenzyme Q), p. 642-656. In F. C. Neidhardt, R. Curtiss, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S.  
15 Reznikoff, M. Riley, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli and Salmonella: cellular and molecular biology*, 2 ed, vol. 1. ASM Press, Washington, D.C.
- Reusch, V. M., Jr. 1984. Lipopolymers, isoprenoids, and the assembly of the gram-positive cell wall. *Crit Rev Microbiol* 11:129-55.
- 20 Wilding, E. I., J. R. Brown, A. P. Bryant, A. F. Chalker, D. J. Holmes, K. A. Ingraham, S. Iordanescu, C. Y. So, M. Rosenberg, and M. N. Gwynn. 2000. Identification, evolution, and essentiality of the mevalonate pathway for isopentenyl diphosphate biosynthesis in gram-positive cocci. *J Bacteriol* 182:4319-27.
- 25 Wilding, E. I., D. Y. Kim, A. P. Bryant, M. N. Gwynn, R. D. Lunsford, D. McDevitt, J. E. Myers, Jr., M. Rosenberg, D. Sylvester, C. V. Stauffacher, and V. W. Rodwell. 2000. Essentiality, expression, and characterization of the class II 3-hydroxy-3-methylglutaryl coenzyme A reductase of *Staphylococcus aureus*. *J Bacteriol* 182:5147-52.
- 30 Beach, M. J., and V. W. Rodwell. 1989. Cloning, sequencing, and overexpression of *mvaA*, which encodes *Pseudomonas mevalonii* 3-hydroxy-3-methylglutaryl coenzyme A reductase. *J Bacteriol* 171:2994-3001.

- Goldstein, J. L., and M. S. Brown. 1990. Regulation of the mevalonate pathway. *Nature* 343:425-30.
- Horbach, S., H. Sahm, and R. Welle. 1993. Isoprenoid biosynthesis in bacteria: two  
5 different pathways? *FEMS Microbiol Lett* 111:135-40.
- Jaworek, A., W. Gruber, and H. U. Bergmeyer. 1974. p. 2127-2131. In H. U. Bergmeyer (ed.), *Methods of Enzymatic Analysis*, 2 ed, vol. 4. Academic Press, New York.
- Johnson, E. A., and W. A. Schroeder. 1996. Microbial carotenoids. *Adv Biochem Eng*  
10 *Biotechnol* 53:119-78.
- Lange, B. M., T. Rujan, W. Martin, and R. Croteau. 2000. Isoprenoid biosynthesis: the evolution of two ancient and distinct pathways across genomes. *Proc Natl Acad Sci U S A* 97:13172-7.
- Meganathan, R. 1996. Biosynthesis of the isoprenoid quinones menaquinone  
15 (vitamin K2) and ubiquinone (coenzyme Q), p. 642-656. In F. C. Neidhardt, R. Curtiss, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli and Salmonella: cellular and molecular biology*, 2 ed, vol. 1. ASM Press, Washington, D.C.
- Michihara, A., M. Sawamura, Y. Nara, K. Ikeda, and Y. Yamori. 1997. Purification and characterization of two mevalonate pyrophosphate decarboxylases from  
20 rat liver: a novel molecular species of 37 kDa. *PG. J Biochem (Tokyo)* 122.
- Reusch, V. M., Jr. 1984. Lipopolymers, isoprenoids, and the assembly of the gram-positive cell wall. *Crit Rev Microbiol* 11:129-55.
- Wilding, E. I., J. R. Brown, A. P. Bryant, A. F. Chalker, D. J. Holmes, K. A. Ingraham,  
25 S. Iordanescu, C. Y. So, M. Rosenberg, and M. N. Gwynn. 2000. Identification, evolution, and essentiality of the mevalonate pathway for isopentenyl diphosphate biosynthesis in gram-positive cocci. *J Bacteriol* 182:4319-27.

- Wilding, E. I., D. Y. Kim, A. P. Bryant, M. N. Gwynn, R. D. Lunsford, D. McDevitt, J. E. Myers, Jr., M. Rosenberg, D. Sylvester, C. V. Stauffacher, and V. W. Rodwell. 2000. Essentiality, expression, and characterization of the class II 3-hydroxy-3-methylglutaryl coenzyme A reductase of *Staphylococcus aureus*. J Bacteriol 182:5147-52.
- 5 Beach, M. J., and V. W. Rodwell. 1989. Cloning, sequencing, and overexpression of *mvaA*, which encodes *Pseudomonas mevalonii* 3-hydroxy-3-methylglutaryl coenzyme A reductase. J Bacteriol 171:2994-3001.
- Goldstein, J. L., and M. S. Brown. 1990. Regulation of the mevalonate pathway. Nature 343:425-30.
- 10 Horbach, S., H. Sahm, and R. Welle. 1993. Isoprenoid biosynthesis in bacteria: two different pathways? FEMS Microbiol Lett 111:135-40.
- Johnson, E. A., and W. A. Schroeder. 1996. Microbial carotenoids. Adv Biochem Eng Biotechnol 53:119-78.
- 15 Lange, B. M., T. Rujan, W. Martin, and R. Croteau. 2000. Isoprenoid biosynthesis: the evolution of two ancient and distinct pathways across genomes. Proc Natl Acad Sci U S A 97:13172-7.
- Meganathan, R. 1996. Biosynthesis of the isoprenoid quinones menaquinone (vitamin K<sub>2</sub>) and ubiquinone (coenzyme Q), p. 642-656. In F. C. Neidhardt, R. Curtiss, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli and Salmonella: cellular and molecular biology*, 2 ed, vol. 1. ASM Press, Washington, D.C.
- 20 Reusch, V. M., Jr. 1984. Lipopolymers, isoprenoids, and the assembly of the gram-positive cell wall. Crit Rev Microbiol 11:129-55.
- 25 Sutherlin, A., M. Hedl, B. Sanchez-Neri, J. W. Burgner, 2nd, C. V. Stauffacher, and V. W. Rodwell. 2002. *Enterococcus faecalis* 3-hydroxy-3-methylglutaryl coenzyme A synthase, an enzyme of isopentenyl diphosphate biosynthesis. J Bacteriol 184:4065-70.

- Wilding, E. I., J. R. Brown, A. P. Bryant, A. F. Chalker, D. J. Holmes, K. A. Ingraham, S. Iordanescu, C. Y. So, M. Rosenberg, and M. N. Gwynn. 2000. Identification, evolution, and essentiality of the mevalonate pathway for isopentenyl diphosphate biosynthesis in gram-positive cocci. *J. Bacteriol.* 182:4319-27.
- Balducci, E., Emanuelli, M., Raffaelli, N., Ruggieri, S., Amici, A., Magni, G., Orsomando, G., Polzonetti, V., Natalini, P. 1995. Assay methods for nicotinamide mononucleotide adenylyltransferase of wide applicability. *Anal. Biochem.*, 228: 64-68
- Begley, T. P., Kinsland, C., Mehl, R. A., Osterman, A., and Dorrestein, P. The biosynthesis of nicotinamide adenine dinucleotides in bacteria. 2001. *Vitam. Horm.* 61: 103-119.
- Garavaglia, S., D'Angelo, I., Emanuelli, M., Carnevali, F., Pierella, F., Magni, G., and Rizzi, M. 2002. Structure of human NMN adenylyltransferase. A key nuclear enzyme for NAD homeostasis. *J. Biol. Chem.* 277: 8524-8530.
- Gerdes, S. Y., Scholle, M. D., D'Souza, M., Bernal, A., Baev, M. V., Farrell, M., Kurnasov, O. V., Daugherty, M. D., Mseeh, F., Polanuyer, B. M., Campbell, J. W., Anantha, S., Shatalin, K. Y., Chowdhury, S. A. K., Fonstein, M. Y., and Osterman, A. L. 2002. From genetic footprinting to antimicrobial drug targets: examples in cofactor biosynthetic pathways. *J. Bacteriol.* 184: 4555-4572.
- Hughes, K. T., Ladika, D., Roth, J. R., and Olivera, B. M. 1983. An indispensable gene for NAD biosynthesis in *Salmonella typhimurium*. *J. Bacteriol.* 155: 213-221.
- Magni, G., Amici, A., Emanuelli, M., Raffaelli, N., and Ruggieri, S. 1999. Enzymology of NAD<sup>+</sup> synthesis. *Adv. Enzymol. Relat. Areas Mol. Biol.* 73: 135-182.
- Mehl, R. A., Kinsland, C., and Begley, T. P. 2000. Identification of the *Escherichia coli* nicotinic acid mononucleotide adenylyltransferase gene. *J. Bacteriol.* 182: 4372-4374.

- Olland, A. M., Underwood, K. W., Czerwinski, R. M., Lo, M-Ch., Aulabaugh, A., Bard, J., Stahl, M. L., Somers, W. S., Sullivan, F. X., and Chopra R. 2002. Identification, characterization, and crystal structure of *Bacillus subtilis* nicotinic acid mononucleotide adenylyltransferase. *J. Biol. Chem.*, 277: 3698-3707.
- Zhang, H., Zhou, T., Kurnasov, O., Cheek, S., Grishin, N. V., Osterman, A., 2002. Crystal structures of *E. coli* nicotinate mononucleotide adenylyltransferase and its complex with deamido-NAD. *Structure*, 10: 6979.
- 10 Zhou, T., Kurnasov, O., Tomchick, D.R., Binns, D.D., Grishin, N.V., Marquez, V.E., Osterman, A.L., Zhang, H. 2002. Structure of human nicotinamide/nicotinic acid mononucleotide adenylyltransferase *J. Biol. Chem.* 277: 13148-13154.
- Werner, E., Ziegler, M., Lerner, F., Schweiger, M., and Heinemann, U. 2002. Crystal structure of human nicotinamide mononucleotide adenylyltransferase in complex with NMN. *FEBS Lett.*, 516: 239-244.
- 15 Balducci, E., Emanuelli, M., Raffaelli, N., Ruggieri, S., Amici, A., Magni, G., Orsomando, G., Polzonetti, V., Natalini, P. 1995. Assay methods for nicotinamide mononucleotide adenylyltransferase of wide applicability. *Anal. Biochem.*, 228: 64-68
- 20 Foster, J.W., and Moat, A.G. 1980. Nicotinamide adenine dinucleotide biosynthesis and pyridine nucleotide cycle metabolism in microbial systems. *Microbial. Rev.* 44: 83-105.
- Hughes, K. T., Olivera, B. M., Roth, J. R., and 1988. Structural gene for NAD synthetase in *Salmonella typhimurium*. *J. Bacteriol.* 170: 2113-2120.
- 25 Gerdes, S. Y., Scholle, M. D., D'Souza, M., Bernal, A., Baev, M. V., Farrell, M., Kurnasov, O. V., Daugherty, M. D., Mseeh, F., Polanuyer, B. M., Campbell, J. W., Anantha, S., Shatalin, K. Y., Chowdhury, S. A. K., Fonstein, M. Y., and Osterman, A. L. 2002. From genetic footprinting to antimicrobial drug targets: examples in cofactor biosynthetic pathways. *J. Bacteriol.* 184: 4555-4572.
- 30 Nessi, C., Albertini, A. M., Speranza, M. L., and Galizzi, A. 1995. The outB gene of *Bacillus subtilis* codes for NAD synthetase. *J. Biol. Chem.* 270: 6181-6185

- Rizzi, M., Nessi, C., Mattevi, A., Coda, A., Bolognesi, M., and Galizzi, A. 1996. Crystal structure of NH<sub>3</sub>-dependent NAD<sup>+</sup> synthetase from *Bacillus subtilis*. EMBO J. 15: 5125-5134.
- 5 Schneider, B. L., and Reitzer, L. J. 1998. *Salmonella typhimurium* nit is nadE: defective nitrogen ntilization and ammonia-Dependent NAD synthetase. J. Bacteriol. 180: 4739-4741.
- Willison, J. C., and Tissot, G. 1994. The *Escherichia coli* efg gene and the *Rhodobacter capsulatus* adgA gene code for NH<sub>3</sub>-dependent NAD synthetase. J. Bacteriol. 176: 3400-3402.
- 10 Aeschlimann, J. R., L. D. Dresser, G. W. Kaatz, and M. J. Rybak. 1999. Effects of NorA inhibitors on in vitro antibacterial activities and postantibiotic effects of levofloxacin, ciprofloxacin, and norfloxacin in genetically related strains of *Staphylococcus aureus*. Antimicrob Agents Chemother 43:335-40.
- 15 Ahmed, M., C. M. Borsch, S. S. Taylor, N. Vazquez-Laslop, and A. A. Neyfakh. 1994. A protein that activates expression of a multidrug efflux transporter upon binding the transporter substrates. J Biol Chem 269:28506-13.
- Ahmed, M., L. Lyass, P. N. Markham, S. S. Taylor, N. Vazquez-Laslop, and A. A. Neyfakh. 1995. Two highly similar multidrug transporters of *Bacillus subtilis* whose expression is differentially regulated. J Bacteriol 177:3904-10.
- 20 Fournier, B., R. Aras, and D. C. Hooper. 2000. Expression of the multidrug resistance transporter NorA from *Staphylococcus aureus* is modified by a two-component regulatory system. J Bacteriol 182:664-71.
- Gill, M. J., N. P. Brenwald, and R. Wise. 1999. Identification of an efflux pump gene, pmrA, associated with fluoroquinolone resistance in *Streptococcus pneumoniae*. Antimicrob Agents Chemother 43:187-9.
- 25 Guz, N. R., F. R. Stermitz, J. B. Johnson, T. D. Beeson, S. Willen, J. Hsiang, and K. Lewis. 2001. Flavonolignan and flavone inhibitors of a *Staphylococcus aureus* multidrug resistance pump: structure-activity relationships. J Med Chem 44:261-8.
- 30 Hsieh, P. C., S. A. Siegel, B. Rogers, D. Davis, and K. Lewis. 1998. Bacteria lacking a multidrug pump: a sensitive tool for drug discovery. Proc Natl Acad Sci U S A 95:6602-6.

- Jonas, B. M., B. E. Murray, and G. M. Weinstock. 2001. Characterization of *emeA*, a NorA homolog and multidrug resistance efflux pump, in *Enterococcus faecalis*. *Antimicrob Agents Chemother* 45:3574-9.
- 5 Kaatz, G. W., and S. M. Seo. 1995. Inducible NorA-mediated multidrug resistance in *Staphylococcus aureus*. *Antimicrob Agents Chemother* 39:2650-5.
- Kaatz, G. W., S. M. Seo, and C. A. Ruble. 1993. Efflux-mediated fluoroquinolone resistance in *Staphylococcus aureus*. *Antimicrob Agents Chemother* 37:1086-94.
- 10 Kaatz, G. W., S. M. Seo, and C. A. Ruble. 1991. Mechanisms of fluoroquinolone resistance in *Staphylococcus aureus*. *J Infect Dis* 163:1080-6.
- Klyachko, K. A., S. Schuldiner, and A. A. Neyfakh. 1997. Mutations affecting substrate specificity of the *Bacillus subtilis* multidrug transporter Bmr. *J Bacteriol* 179:2189-93.
- 15 Markham, P. N., and A. A. Neyfakh. 1996. Inhibition of the multidrug transporter NorA prevents emergence of norfloxacin resistance in *Staphylococcus aureus*. *Antimicrob Agents Chemother* 40:2673-4.
- Markham, P. N., E. Westhaus, K. Klyachko, M. E. Johnson, and A. A. Neyfakh. 1999. Multiple novel inhibitors of the NorA multidrug transporter of *Staphylococcus aureus*. *Antimicrob Agents Chemother* 43:2404-8.
- 20 Neyfakh, A. A. 1992. The multidrug efflux transporter of *Bacillus subtilis* is a structural and functional homolog of the *Staphylococcus* NorA protein. *Antimicrob Agents Chemother* 36:484-5.
- Neyfakh, A. A., C. M. Borsch, and G. W. Kaatz. 1993. Fluoroquinolone resistance protein NorA of *Staphylococcus aureus* is a multidrug efflux transporter. *Antimicrob Agents Chemother* 37:128-9.
- 25 Ng, E. Y., M. Trucksis, and D. C. Hooper. 1994. Quinolone resistance mediated by *norA*: physiologic characterization and relationship to *flqB*, a quinolone resistance locus on the *Staphylococcus aureus* chromosome. *Antimicrob Agents Chemother* 38:1345-55.
- 30 Paulsen, I. T., M. H. Brown, and R. A. Skurray. 1996. Proton-dependent multidrug efflux systems. *Microbiol Rev* 60:575-608.



- Putman, M., H. W. van Veen, B. Poolman, and W. N. Konings. 1999. Restrictive use of detergents in the functional reconstitution of the secondary multidrug transporter LmrP. *Biochemistry* 38:1002-8.
- 5 Yoshida, H., M. Bogaki, S. Nakamura, K. Ubukata, and M. Konno. 1990. Nucleotide sequence and characterization of the *Staphylococcus aureus* norA gene, which confers resistance to quinolones. *J Bacteriol* 172:6942-9.
- Yu, J. L., L. Grinius, and D. C. Hooper. 2002. NorA functions as a multidrug efflux protein in both cytoplasmic membrane vesicles and reconstituted
- 10 proteoliposomes. *J Bacteriol* 184:1370-7.
- Britton, R. A., B. S. Powell, D. L. Court, and J. R. Lupski. 1997. Characterization of mutations affecting the *Escherichia coli* essential GTPase era that suppress two temperature-sensitive dnaG alleles. *J Bacteriol* 179:4575-82.
- Chen, S. M., H. E. Takiff, A. M. Barber, G. C. Dubois, J. C. Bardwell, and D. L. Court.
- 15 1990. Expression and characterization of RNase III and Era proteins. Products of the *rnc* operon of *Escherichia coli*. *J Biol Chem* 265:2888-95.
- Inada, T., K. Kawakami, S. M. Chen, H. E. Takiff, D. L. Court, and Y. Nakamura. 1989. Temperature-sensitive lethal mutant of era, a G protein in *Escherichia coli*. *J Bacteriol* 171:5017-24.
- 20 Kok, J., K. A. Trach, and J. A. Hoch. 1994. Effects on *Bacillus subtilis* of a conditional lethal mutation in the essential GTP-binding protein Obg. *J Bacteriol* 176:7155-60.
- Lanzetta, P. A., L. J. Alvarez, P. S. Reinach, and O. A. Candia. 1979. An improved assay for nanomole amounts of inorganic phosphate. *Anal Biochem* 100:95-7.
- 25 Lin, B., J. M. Skidmore, A. Bhatt, S. M. Pfeffer, L. Pawloski, and J. R. Maddock. 2001. Alanine scan mutagenesis of the switch I domain of the *Caulobacter crescentus* CgtA protein reveals critical amino acids required for in vivo function. *Mol Microbiol* 39:924-34.
- March, P. E. 1992. Membrane-associated GTPases in bacteria. *Mol Microbiol*
- 30 6:1253-7.
- Scott, J. M., and W. G. Haldenwang. 1999. Obg, an essential GTP binding protein of *Bacillus subtilis*, is necessary for stress activation of transcription factor sigma(B). *J Bacteriol* 181:4653-60.

- Scott, J. M., J. Ju, T. Mitchell, and W. G. Haldenwang. 2000. The *Bacillus subtilis* GTP binding protein obg and regulators of the sigma(B) stress response transcription factor cofractionate with ribosomes. *J Bacteriol* 182:2771-7.
- 5 Sood, P., C. G. Lerner, T. Shimamoto, Q. Lu, and M. Inouye. 1994. Characterization of the autophosphorylation of Era, an essential *Escherichia coli* GTPase. *Mol Microbiol* 12:201-8.
- Trach, K., and J. A. Hoch. 1989. The *Bacillus subtilis* spo0B stage 0 sporulation operon encodes an essential GTP-binding protein. *J Bacteriol* 171:1362-71.
- 10 Welsh, K. M., K. A. Trach, C. Folger, and J. A. Hoch. 1994. Biochemical characterization of the essential GTP-binding protein Obg of *Bacillus subtilis*. *J Bacteriol* 176:7161-8.
- Hwang, J. and Inouye, M. An essential GTPase, Der, containing double GTP-binding domains from *Escherichia coli* and *Thermotoga maritima*. *Journal of Biological Chemistry*. 33:31415-31421. 2001.
- 15 Britton, R. A., B. S. Powell, D. L. Court, and J. R. Lupski. 1997. Characterization of mutations affecting the *Escherichia coli* essential GTPase era that suppress two temperature-sensitive dnaG alleles. *J Bacteriol* 179:4575-82.
- Chen, S. M., H. E. Takiff, A. M. Barber, G. C. Dubois, J. C. Bardwell, and D. L. Court. 20 1990. Expression and characterization of RNase III and Era proteins. Products of the mrc operon of *Escherichia coli*. *J Biol Chem* 265:2888-95.
- Sood, P., C. G. Lerner, T. Shimamoto, Q. Lu, and M. Inouye. 1994. Characterization of the autophosphorylation of Era, an essential *Escherichia coli* GTPase. *Mol Microbiol* 12:201-8.

25

**WHAT IS CLAIMED IS:**

- 5 1. A purified or isolated *Alloioococcus otitidis* nucleic acid sequence comprising a  
nucleotide sequence selected from one of odd numbered sequences set forth  
in Seq. ID Nos: 1 to Seq. ID Nos: 105, wherein expression of said nucleic acid  
is essential for the proliferation of a cell.
- 10 2. A purified or isolated nucleic acid of *Alloioococcus otitidis* comprising a  
fragment of one of odd numbered sequences set forth in Seq. ID Nos: 1 to  
Seq. ID Nos: 105 said fragment selected from the group consisting of  
fragments comprising at least 10, at least 20, at least 25, at least 30, at least  
50 and more than 50 consecutive nucleotides of one of one of odd numbered  
15 sequences set forth in Seq. ID Nos: 1 to Seq. ID Nos: 105.
3. A purified or isolated antisense nucleic acid comprising a nucleotide  
sequence complementary to at least a portion of an intragenic sequence,  
intergenic sequence, sequences spanning at least a portion of two or more  
20 genes, 5' noncoding region, or 3' noncoding region within an operon  
comprising a proliferation-required gene of *Alloioococcus otitidis* whose activity  
or expression is inhibited by an antisense nucleic acid and selected from one  
of odd numbered sequences set forth in Seq. ID Nos: 1 to Seq. ID Nos: 105.
- 25 4. A purified or isolated nucleic acid comprising a nucleotide sequence having at  
least 70% identity to a nucleotide sequence selected from one of odd  
numbered sequences set forth in Seq. ID Nos: 1 to Seq. ID Nos: 105,  
fragments comprising at least 25 consecutive nucleotides selected from one  
of odd numbered sequences set forth in Seq. ID Nos: 1 to Seq. ID Nos: 105,  
30 the nucleotide sequences complementary to one of odd numbered sequences  
set forth in Seq. ID Nos: 1 to Seq. ID Nos: 105, and the sequences  
complementary to fragments comprising at least 25 consecutive nucleotides

of one of odd numbered sequences set forth in Seq. ID Nos: 1 to Seq. ID Nos: 105.

- 5 5. A vector comprising a promoter operably linked to a nucleic acid encoding a polypeptide whose expression is inhibited by an antisense nucleic acid comprising a nucleotide sequence of any one of odd numbered sequences set forth in Seq. ID Nos: 1 to Seq. ID Nos: 105.
- 10 5. A purified or isolated polypeptide of *Alloicoccus otitidis* comprising a polypeptide whose expression is inhibited by an antisense nucleic acid comprising a nucleotide sequence of one of odd numbered sequences set forth in Seq. ID Nos: 1 to Seq. ID Nos: 105, or a fragment selected from the group consisting of fragments comprising at least 5, at least 10, at least 20, at least 30, at least 40, at least 50, at least 60 or more than 60 consecutive amino acids of one of the said polypeptides.
- 15 6. A purified or isolated *Alloicoccus otitidis* polypeptide comprising a amino acid sequence having at least 25% amino acid identity to a polypeptide whose expression is inhibited by a nucleic acid comprising a nucleotide sequence selected from one of odd numbered sequences set forth in Seq. ID Nos: 1 to Seq. ID Nos: 105, or at least 25% amino acid identity to a fragment comprising at least 10, at least 20, at least 30, at least 40, at least 50, at least 60 or more than 60 consecutive amino acids of a polypeptide whose expression is inhibited by a nucleic acid comprising a nucleotide sequence selected from the group consisting of one of odd numbered sequences set forth in Seq. ID Nos: 1 to Seq. ID Nos: 105.
- 20 25 7. A purified or isolated *Alloicoccus otitidis* polypeptide comprising selected from one of the even numbered sequences set forth in Seq. ID Nos: 2 to Seq. ID Nos: 106, wherein the polypeptide is essential for the proliferation of a cell.
- 30

8. A method of producing an *Alloiococcus otitidis* polypeptide comprising introducing into a cell a vector comprising a promoter operably linked to a nucleic acid comprising a nucleotide sequence encoding a polypeptide whose expression is essential for the proliferation and viability of *Alloiococcus otitidis*, and which is inhibited by an antisense nucleic acid, and which is selected from one of odd numbered sequences set forth in Seq. ID Nos: 1 to Seq. ID Nos: 105.
9. A method of inhibiting the proliferation of *Alloiococcus otitidis* in an individual comprising inhibiting the activity or reducing the amount of a gene product whose expression is inhibited by an antisense nucleic acid comprising a nucleotide sequence selected from one of odd numbered sequences set forth in Seq. ID Nos: 1 to Seq. ID Nos: 105 or inhibiting the activity or reducing the amount of a nucleic acid encoding said gene product.
10. A method for identifying a compound which influences the activity of an *Alloiococcus otitidis* gene product, which is required for proliferation, said gene product comprising a gene product whose expression is inhibited by an antisense nucleic acid comprising a nucleotide sequence selected from one of odd numbered sequences set forth in Seq. ID Nos: 1 to Seq. ID Nos: 105, said method comprising:
- (a) contacting said gene product with a candidate compound; and
  - (b) determining whether said compound influences the activity of said gene product.
11. A method for identifying a compound or an antisense nucleic acid having the ability to reduce activity or level of a *Alloiococcus otitidis* gene product, which is required for proliferation, said gene product comprising a gene product whose activity or expression is inhibited by an antisense nucleic acid

comprising a nucleotide sequence selected from one of odd numbered sequences set forth in Seq. ID Nos: 1 to Seq. ID Nos: 105, said method comprising the steps of:

- (a) contacting a target gene or RNA encoding said gene product with a candidate compound or antisense nucleic acid; and
- (b) measuring the activity of said target.

5

10

15

20

25

30

13. A method for inhibiting cellular proliferation of *Alloicoccus otitidis* comprising introducing an effective amount of a compound with activity against a gene whose activity or expression is essential for cellular proliferation, and which is inhibited by an antisense nucleic acid comprising a nucleotide sequence selected from one of odd numbered sequences set forth in Seq. ID Nos: 1 to Seq. ID Nos: 105, or a compound with activity against the product of said gene into a population of *Alloicoccus otitidis* cells expressing said gene.
13. A composition comprising an effective concentration of an antisense nucleic acid comprising a nucleotide sequence selected from one of odd numbered sequences set forth in Seq. ID Nos: 1 to Seq. ID Nos: 105, or a proliferation-inhibiting portion thereof in a pharmaceutically acceptable carrier.
14. A method for identifying a compound having the ability to inhibit proliferation of *Alloicoccus otitidis* cell comprising:
  - (a) identifying a homologue of a gene or gene product whose activity or level is inhibited by a nucleic acid comprising a nucleotide sequence selected from one of odd numbered sequences set forth in Seq. ID Nos: 1 to Seq. ID Nos: 105, in a test cell, wherein said test cell is not *Alloicoccus otitidis*;
  - (a) identifying an inhibitory nucleic acid sequence which inhibits the activity of said homologue in said test cell;
  - (b) contacting said test cell with a sublethal level of said inhibitory nucleic acid, thus sensitizing said cell;
  - (c) contacting the sensitized cell of step (c) with a compound; and

- (d) determining the degree to which said compound inhibits proliferation of said sensitized cell relative to a cell which does not contain said inhibitory nucleic acid.

- 5 16. A method for identifying a compound having activity against a biological pathway required for proliferation comprising:
- 10 (a) sensitizing a cell by providing a sublethal level of an antisense nucleic acid complementary to a nucleic acid encoding a gene product required for proliferation, wherein the activity or expression of said gene product is inhibited by an antisense nucleic acid comprising a nucleotide sequence selected from one of odd numbered sequences set forth in Seq. ID Nos: 1 to Seq. ID Nos: 105, in said cell to reduce the activity or amount of said gene product;
- 15 (a) contacting the sensitized cell with a compound; and  
(b) determining the degree to which said compound inhibits the growth of said sensitized cell relative to a cell which does not contain said antisense nucleic acid.
- 20 17. A method for identifying a compound having the ability to inhibit one of the *Alloiococcus otitidis* polypeptides encoded by a polynucleotide selected from one of odd numbered sequences set forth in Seq. ID Nos: 1 to Seq. ID Nos: 105, and which is essential for cellular proliferation comprising:
- 25 (a) contacting a cell which expresses the polypeptide with the compound; and  
(b) determining whether said compound reduces proliferation of said contacted cell by acting on said gene product.
- 30 18. A method for identifying a compound having the ability to inhibit one of the purified and isolated *Alloiococcus otitidis* polypeptides selected from one of the even numbered sequences set forth in Seq. ID No.: 2 to Seq. ID No.: 106, and which is essential for cellular proliferation comprising:

- (c) contacting the purified and isolated polypeptide with the compound *in vitro* in the presence or absence of a substrate, which is essential for the activity of the polypeptide; and
- (d) determining the effect of the compound on the polypeptide by measuring the effect of the polypeptide on the substrate.

19. A compound which interacts with an *Alloiococcus otitidis* polypeptide selected from one of the even numbered sequences set forth in Seq. ID No.: 2 to Seq. ID No.: 106 and inhibits its activity.
20. A method for manufacturing an antimicrobial compound comprising the steps of screening one or more candidate compounds to identify a compound that reduces the activity or level of an *Alloiococcus otitidis* polypeptide selected from one of the even numbered sequences set forth in Seq. ID No.: 2 to Seq. ID No.: 106, said polypeptide comprising a gene product whose activity or expression is inhibited by an antisense nucleic acid comprising a nucleotide sequence selected from one of the odd numbered sequences set forth in Seq. ID No.: 1 to Seq. ID No. 105; and manufacturing the compound so identified.
21. A compound which inhibits proliferation of *Alloiococcus otitidis* by interacting with a gene encoding a polypeptide that is required for proliferation or with a polypeptide required for proliferation, wherein said polypeptide is selected from the group consisting of a gene product having at least 70% nucleotide sequence identity from one of the odd numbered sequences set forth in Seq. ID No.: 1 to Seq. ID No. 105, polypeptide encoded by a nucleic acid having at least 70% nucleotide sequence identity to a nucleic acid encoding a polypeptide whose expression is inhibited by an antisense nucleic acid comprising a nucleotide sequence selected from one of the odd numbered sequences set forth in Seq. ID No.: 1 to Seq. ID No. 105, a polypeptide having at least 25% amino acid identity to a gene product whose expression is inhibited by an antisense nucleic acid comprising a nucleotide sequence selected one of the odd numbered sequences set forth in Seq. ID No.: 1 to



Seq. ID No. 105, a polypeptide encoded by a nucleic acid comprising a nucleotide sequence which hybridizes to a nucleic acid selected from one of the odd numbered sequences set forth in Seq. ID No.: 1 to Seq. ID No. 105 under stringent conditions, a gene product encoded by a nucleic acid  
5 comprising a nucleotide sequence which hybridizes to a nucleic acid selected from one of the odd numbered sequences set forth in Seq. ID No.: 1 to Seq. ID No. 105 under moderate conditions, and a gene product whose activity may be complemented by the gene product whose activity is inhibited by a nucleic acid selected from one of the odd numbered sequences set forth in  
10 Seq. ID No.: 1 to Seq. ID No. 105.

## SEQUENCE LISTING

<110> American Cyanamid Company, and Murphy, Ellen and Projan, Stephen, j.

<120> Alloiococcus otitidis Infectious Disease Targets

<130> Application 1

<160> 106

<170> PatentIn version 3.1

<210> 1

<211> 426

<212> DNA

<213> Alloiococcus otitidis

<220>

<221> CDS

<222> (73)..(426)

<223>

<400> 1

aagacaaaaa agaagaggga aaagatctta agacacttcc ctaagtctga acatattcta 60

ggagggttac aa gtg att aca gga atg ggt gtg gat att gtt gaa atg agc 111  
Met Ile Thr Gly Met Gly Val Asp Ile Val Glu Met Ser  
1 5 10

cgg att caa gct gtt tgg gac cga aag ccc agc ttt gcc cag cgg att 159  
Arg Ile Gln Ala Val Trp Asp Arg Lys Pro Ser Phe Ala Gln Arg Ile  
15 20 25

tta acc caa agg gag ttg gct tat ttc gag aaa gcg act ggt agg cgg 207  
Leu Thr Gln Arg Glu Leu Ala Tyr Phe Glu Lys Ala Thr Gly Arg Arg  
30 35 40 45

aga att gaa ttc cta gcg gga cgg ttt gcc ggt aaa gaa gct tac agt 255  
Arg Ile Glu Phe Leu Ala Gly Arg Phe Ala Gly Lys Glu Ala Tyr Ser  
50 55 60

aaa gcc ttg gga act ggt att gga cgc ttg agc ttt aaa gat att gaa 303  
Lys Ala Leu Gly Thr Gly Ile Gly Arg Leu Ser Phe Lys Asp Ile Glu  
65 70 75

atc cta atc aat gac caa ggc cag cca gtc cta aca tct cat cct aaa 351  
Ile Leu Ile Asn Asp Gln Gly Gln Pro Val Leu Thr Ser His Pro Lys  
80 85 90

gct gcc cgg gcc ttg att tca att tct cac act aga gac ctc tgc ctg 399  
Ala Gly Arg Ala Leu Ile Ser Ile Ser His Thr Arg Asp Leu Cys Leu  
95 100 105

gcc cag gtc ctt tta cag gaa aat tga 426  
Ala Gln Val Leu Leu Gln Glu Asn  
110 115

<210> 2  
 <211> 117  
 <212> PRT  
 <213> *Alloiococcus otitidis*

<400> 2  
 Met Ile Thr Gly Met Gly Val Asp Ile Val Glu Met Ser Arg Ile Gln  
 1 5 10 15

Ala Val Trp Asp Arg Lys Pro Ser Phe Ala Gln Arg Ile Leu Thr Gln  
 20 25 30

Arg Glu Leu Ala Tyr Phe Glu Lys Ala Thr Gly Arg Arg Arg Ile Glu  
 35 40 45

Phe Leu Ala Gly Arg Phe Ala Gly Lys Glu Ala Tyr Ser Lys Ala Leu  
 50 55 60

Gly Thr Gly Ile Gly Arg Leu Ser Phe Lys Asp Ile Glu Ile Leu Ile  
 65 70 75 80

Asn Asp Gln Gly Gln Pro Val Leu Thr Ser His Pro Lys Ala Gly Arg  
 85 90 95

Ala Leu Ile Ser Ile Ser His Thr Arg Asp Leu Cys Leu Ala Gln Val  
 100 105 110

Leu Leu Gln Glu Asn  
 115

<210> 3  
 <211> 1410  
 <212> DNA  
 <213> *Alloiococcus otitidis*

<220>  
 <221> CDS  
 <222> (16)..(1410)  
 <223>

<400> 3  
 ataggagtca taacg gtg tct tgg aaa tta aaa gag att gcc cag gca gtt. 51  
 Met Ser Trp Lys Leu Lys Glu Ile Ala Gln Ala Val  
 1 5 10

ggg gga gag cta gtt agt gcg gac ggc cag gag gag gtc acc ggg gtc 99  
 Gly Gly Glu Leu Val Ser Ala Asp Gly Gln Glu Glu Val Thr Gly Val

15	20	25	
cac ttt gat tca agg cga ctt gaa cca ggt gac ttg ttt gtt cct att			147
His Phe Asp Ser Arg Arg Leu Glu Pro Gly Asp Leu Phe Val Pro Ile			
30	35	40	
tta ggc cag cgg gat ggt cat gat ttt gcc caa gcc gcc cta gac caa			195
Leu Gly Gln Arg Asp Gly His Asp Phe Ala Gln Ala Ala Leu Asp Gln			
45	50	55	60
gga gct agc gga gcc ttt tgg gcc aaa gat tca agc tta gcc cct aaa			243
Gly Ala Ser Gly Ala Phe Trp Ala Lys Asp Ser Ser Leu Ala Pro Lys			
	65	70	75
ggt ctt ccc ttg atc aag gta gaa gat agc tac cag gcc cta gtt gac			291
Gly Leu Pro Leu Ile Lys Val Glu Asp Ser Tyr Gln Ala Leu Val Asp			
	80	85	90
ctg gcc aag tgg cat ctt gaa gct gtc gca cct atg aaa att gcc atc			339
Leu Ala Lys Trp His Leu Glu Ala Val Ala Pro Met Lys Ile Ala Ile			
	95	100	105
acc ggc agt aat ggg aag acc act act aag gac atg gtg gct agt gtg			387
Thr Gly Ser Asn Gly Lys Thr Thr Thr Lys Asp Met Val Ala Ser Val			
	110	115	120
gtg ggc caa gca ttt aag tgt cac aaa aca gtt agc aac tta aat aat			435
Val Gly Gln Ala Phe Lys Cys His Lys Thr Val Ser Asn Leu Asn Asn			
125	130	135	140
gaa ctt ggc gtg ccc atg act atc tta gct atg cct gca gac tgc cag			483
Glu Leu Gly Val Pro Met Thr Ile Leu Ala Met Pro Ala Asp Cys Gln			
	145	150	155
gtc ata gtt gtt gaa atg ggc atg gat gga cca ggt cag atc tcg gcc			531
Val Ile Val Val Glu Met Gly Met Asp Gly Pro Gly Gln Ile Ser Ala			
	160	165	170
ttg tcc aaa ctc ttg cag cct gac att gcc att atc acc atg att ggc			579
Leu Ser Lys Leu Leu Gln Pro Asp Ile Ala Ile Ile Thr Met Ile Gly			
	175	180	185
gag gcc cac atc gag ttc ttt ggg tca agg gac aaa att gcc cag gcc			627
Glu Ala His Ile Glu Phe Phe Gly Ser Arg Asp Lys Ile Ala Gln Ala			
	190	195	200
aaa ctg gaa att cta gat ggc cta agc gac cag ggc gtc ttt att gcc			675
Lys Leu Glu Ile Leu Asp Gly Leu Ser Asp Gln Gly Val Phe Ile Ala			
205	210	215	220
aac ggg gat gaa ccc ctg ctt gag tct gcc ttg aac cac cac ccc cac			723
Asn Gly Asp Glu Pro Leu Leu Glu Ser Ala Leu Asn His His Pro His			
	225	230	235
agc ctg cgt ttt ggc caa tcg ccc cac aat gac att tat cct ttg acc			771
Ser Leu Arg Phe Gly Gln Ser Pro His Asn Asp Ile Tyr Pro Leu Thr			
	240	245	250

act gag att gga cag cgg caa agc cag ttc acc ctt aac ctg gac cct Thr Glu Ile Gly Gln Arg Gln Ser Gln Phe Thr Leu Asn Leu Asp Pro 255 260 265	819
agt ctg caa ttt acc atc cct tca cca gga aaa tat aat gtc att aac Ser Leu Gln Phe Thr Ile Pro Ser Pro Gly Lys Tyr Asn Val Ile Asn 270 275 280	867
gcc cta gct gca gtc ttg gta gcc cag gtc ctg gac ttg gac ctc caa Ala Leu Ala Ala Val Leu Val Ala Gln Val Leu Asp Leu Asp Leu Gln 285 290 295 300	915
cta gct gtc cag ggc ttg gcc cag ttt cag cta agc aaa aac cgg ctg Leu Ala Val Gln Gly Leu Ala Gln Phe Gln Leu Ser Lys Asn Arg Leu 305 310 315	963
gaa tgg cta aaa ggc tat aag cag gcc cac tta tta aat gat gct tac Glu Trp Leu Lys Gly Tyr Lys Gln Ala His Leu Leu Asn Asp Ala Tyr 320 325 330	1011
aat gct agt ccc act tcc atg aag gcg gtc ttg gat tat ttc agc cat Asn Ala Ser Pro Thr Ser Met Lys Ala Val Leu Asp Tyr Phe Ser His 335 340 345	1059
ttg gac cta gat ggg gag aag ata gcg gtt tta ggg gac ttg cgg gag Leu Asp Leu Asp Gly Glu Lys Ile Ala Val Leu Gly Asp Leu Arg Glu 350 355 360	1107
tta ggg tct ttg tcc ggt caa ctc cac cgg tca ctt agt caa gcc atc Leu Gly Ser Leu Ser Gly Gln Leu His Arg Ser Leu Ser Gln Ala Ile 365 370 375 380	1155
gac ccc aaa ctt tta gac cgg gtt gtc tta tat gga cca gaa atg gca Asp Pro Lys Leu Leu Asp Arg Val Val Leu Tyr Gly Pro Glu Met Ala 385 390 395	1203
gcc ctc tac cag gtc ttg aag gct gat ttt gat cct gac cac ttg act Ala Leu Tyr Gln Val Leu Lys Ala Asp Phe Asp Pro Asp His Leu Thr 400 405 410	1251
tat ttc cca gag gat cga aaa gcc ttg acc gac ttt tta aaa gaa atc Tyr Phe Pro Glu Asp Arg Lys Ala Leu Thr Asp Phe Leu Lys Glu Ile 415 420 425	1299
atg ggc cca tct tct tat ctt ttg ttg aag tcc agt cta gga aca ggt Met Gly Pro Ser Ser Tyr Leu Leu Leu Lys Ser Ser Leu Gly Thr Gly 430 435 440	1347
ctg ctt gaa gtg gtc caa gcc cta agt caa aaa gaa gat gat gaa aac Leu Leu Glu Val Val Gln Ala Leu Ser Gln Lys Glu Asp Asp Glu Asn 445 450 455 460	1395
cag ccc ctg gac taa Gln Pro Leu Asp	1410

&lt;210&gt; 4

&lt;211&gt; 464

&lt;212&gt; PRT

<213> *Alloiococcus otitidis*

&lt;400&gt; 4

Met Ser Trp Lys Leu Lys Glu Ile Ala Gln Ala Val Gly Gly Glu Leu  
1 5 10 15

Val Ser Ala Asp Gly Gln Glu Glu Val Thr Gly Val His Phe Asp Ser  
20 25 30

Arg Arg Leu Glu Pro Gly Asp Leu Phe Val Pro Ile Leu Gly Gln Arg  
35 40 45

Asp Gly His Asp Phe Ala Gln Ala Ala Leu Asp Gln Gly Ala Ser Gly  
50 55 60

Ala Phe Trp Ala Lys Asp Ser Ser Leu Ala Pro Lys Gly Leu Pro Leu  
65 70 75 80

Ile Lys Val Glu Asp Ser Tyr Gln Ala Leu Val Asp Leu Ala Lys Trp  
85 90 95

His Leu Glu Ala Val Ala Pro Met Lys Ile Ala Ile Thr Gly Ser Asn  
100 105 110

Gly Lys Thr Thr Thr Lys Asp Met Val Ala Ser Val Val Gly Gln Ala  
115 120 125

Phe Lys Cys His Lys Thr Val Ser Asn Leu Asn Asn Glu Leu Gly Val  
130 135 140

Pro Met Thr Ile Leu Ala Met Pro Ala Asp Cys Gln Val Ile Val Val  
145 150 155 160

Glu Met Gly Met Asp Gly Pro Gly Gln Ile Ser Ala Leu Ser Lys Leu  
165 170 175

Leu Gln Pro Asp Ile Ala Ile Ile Thr Met Ile Gly Glu Ala His Ile  
180 185 190

Glu Phe Phe Gly Ser Arg Asp Lys Ile Ala Gln Ala Lys Leu Glu Ile  
195 200 205

Leu Asp Gly Leu Ser Asp Gln Gly Val Phe Ile Ala Asn Gly Asp Glu  
210 215 220

Pro Leu Leu Glu Ser Ala Leu Asn His His Pro His Ser Leu Arg Phe  
225 230 235 240

Gly Gln Ser Pro His Asn Asp Ile Tyr Pro Leu Thr Thr Glu Ile Gly  
245 250 255

Gln Arg Gln Ser Gln Phe Thr Leu Asn Leu Asp Pro Ser Leu Gln Phe  
260 265 270

Thr Ile Pro Ser Pro Gly Lys Tyr Asn Val Ile Asn Ala Leu Ala Ala  
275 280 285

Val Leu Val Ala Gln Val Leu Asp Leu Asp Leu Gln Leu Ala Val Gln  
290 295 300

Gly Leu Ala Gln Phe Gln Leu Ser Lys Asn Arg Leu Glu Trp Leu Lys  
305 310 315 320

Gly Tyr Lys Gln Ala His Leu Leu Asn Asp Ala Tyr Asn Ala Ser Pro  
325 330 335

Thr Ser Met Lys Ala Val Leu Asp Tyr Phe Ser His Leu Asp Leu Asp  
340 345 350

Gly Glu Lys Ile Ala Val Leu Gly Asp Leu Arg Glu Leu Gly Ser Leu  
355 360 365

Ser Gly Gln Leu His Arg Ser Leu Ser Gln Ala Ile Asp Pro Lys Leu  
370 375 380

Leu Asp Arg Val Val Leu Tyr Gly Pro Glu Met Ala Ala Leu Tyr Gln  
385 390 395 400

Val Leu Lys Ala Asp Phe Asp Pro Asp His Leu Thr Tyr Phe Pro Glu  
405 410 415

Asp Arg Lys Ala Leu Thr Asp Phe Leu Lys Glu Ile Met Gly Pro Ser  
420 425 430

Ser Tyr Leu Leu Leu Lys Ser Ser Leu Gly Thr Gly Leu Leu Glu Val  
 435 440 445

Val Gln Ala Leu Ser Gln Lys Glu Asp Asp Glu Asn Gln Pro Leu Asp  
 450 455 460

<210> 5

<211> 1284

<212> DNA

<213> *Alloioicoccus otitidis*

<220>

<221> CDS

<222> (7)..(1284)

<223>

<400> 5

gattgg atg aat ata atg aaa aaa cta atc atc aac ggt ggc cgg acc 48  
 Met Asn Ile Met Lys Lys Leu Ile Ile Asn Gly Gly Arg Thr  
 1 5 10

ctc aag ggt gaa gtc acg gta tca ggg gcc aaa aat agt acg gtg gct 96  
 Leu Lys Gly Glu Val Thr Val Ser Gly Ala Lys Asn Ser Thr Val Ala  
 15 20 25 30

ctc att cca gca tct att tta gca gac agc ccg gta atc cta gag ggg 144  
 Leu Ile Pro Ala Ser Ile Leu Ala Asp Ser Pro Val Ile Leu Glu Gly  
 35 40 45

gta ccc gat atc cag gat gtt cat tcc cta ctg gag att tta aat gaa 192  
 Val Pro Asp Ile Gln Asp Val His Ser Leu Leu Glu Ile Leu Asn Glu  
 50 55 60

atg aat gtc aag acc gac ttt gac gga aac act ttg acc att gac cca 240  
 Met Asn Val Lys Thr Asp Phe Asp Gly Asn Thr Leu Thr Ile Asp Pro  
 65 70 75

aga gaa atg gtc tct atc ccc atg cca agt ggt aag atc caa agc ttg 288  
 Arg Glu Met Val Ser Ile Pro Met Pro Ser Gly Lys Ile Gln Ser Leu  
 80 85 90

cgg gct tcc tac tac ttt atg gga gcc ctc ttg gcc aaa ttc ggt aaa 336  
 Arg Ala Ser Tyr Tyr Phe Met Gly Ala Leu Leu Ala Lys Phe Gly Lys  
 95 100 105 110

ggg gta gtc ggt ctt ccc ggt ggt tgc ttc ctg ggg cca cga ccc atc 384  
 Gly Val Val Gly Leu Pro Gly Gly Cys Phe Leu Gly Pro Arg Pro Ile  
 115 120 125

gac caa cac ttg aaa ggc ttc cgc ctg ctt gga gca gat gtg gat aat 432  
 Asp Gln His Leu Lys Gly Phe Arg Leu Leu Gly Ala Asp Val Asp Asn  
 130 135 140



gaa atg ggg gcc atg tac ctt aaa acc agt gat tca ggc cta gtg ggt Glu Met Gly Ala Met Tyr Leu Lys Thr Ser Asp Ser Gly Leu Val Gly 145 150 155	480
agt cgg att tac tta gat gtt gtt tcg att ggt gca acc att aat atc Ser Arg Ile Tyr Leu Asp Val Val Ser Ile Gly Ala Thr Ile Asn Ile 160 165 170	528
atg tta gcc gct gtt agg gcc caa ggt cgg acg gtc att gag aat gcg Met Leu Ala Ala Val Arg Ala Gln Gly Arg Thr Val Ile Glu Asn Ala 175 180 185 190	576
gcc cga gaa cca gaa att att gat gtt gcc acc ctc ttg aac aag atg Ala Arg Glu Pro Glu Ile Ile Asp Val Ala Thr Leu Leu Asn Lys Met 195 200 205	624
ggg gct aaa ata cgt ggg gct ggc act gat atg atc cgg att gaa ggg Gly Ala Lys Ile Arg Gly Ala Gly Thr Asp Met Ile Arg Ile Glu Gly 210 215 220	672
gtt gac cag ctg act ggc tgc cag cac tcc atc atc ccc gac cgg att Val Asp Gln Leu Thr Gly Cys Gln His Ser Ile Ile Pro Asp Arg Ile 225 230 235	720
gaa gct ggg acc tac ctg gct att gca gcg gca gct ggg gag gat gtc Glu Ala Gly Thr Tyr Leu Ala Ile Ala Ala Ala Gly Glu Asp Val 240 245 250	768
ctg gta aac aat gtt ata gtt gaa cat att gat agt tta att gcc aaa Leu Val Asn Asn Val Ile Val Glu His Ile Asp Ser Leu Ile Ala Lys 255 260 265 270	816
ctc gac gaa att ggt att gac ctg gac atc ggc gaa gac agt atc cgg Leu Asp Glu Ile Gly Ile Asp Leu Asp Ile Gly Glu Asp Ser Ile Arg 275 280 285	864
gtg aaa gcc ccc agt aaa cct ttg cag cct gtt acc atc aaa acc ctg Val Lys Ala Pro Ser Lys Pro Leu Gln Pro Val Thr Ile Lys Thr Leu 290 295 300	912
cct tac cct ggt ttt gcc act gac ctc cag cag ccc atc acc cct ctc Pro Tyr Pro Gly Phe Ala Thr Asp Leu Gln Gln Pro Ile Thr Pro Leu 305 310 315	960
ttg ctt ctg gcc aaa ggg gag tcc gtt atc acc gat acc atc tat cct Leu Leu Leu Ala Lys Gly Glu Ser Val Ile Thr Asp Thr Ile Tyr Pro 320 325 330	1008
aaa cgg gtt aag cac atc cct gag ctg gaa cgg atg ggg gcc aat atc Lys Arg Val Lys His Ile Pro Glu Leu Glu Arg Met Gly Ala Asn Ile 335 340 345 350	1056
cgg gtc gaa agc gat atc atc ctc att gaa ggt ggc cac ccc ctc aag Arg Val Glu Ser Asp Ile Ile Leu Ile Glu Gly Gly His Pro Leu Lys 355 360 365	1104

ggg gca gaa gtg gaa gcc agt gat tta aga gcc ggg gct tgc ttg att 1152  
 Gly Ala Glu Val Glu Ala Ser Asp Leu Arg Ala Gly Ala Cys Leu Ile  
                   370                  375                  380

aat gca ggt ttg atc gcg gaa ggt cag acg gaa att act ggc gtt gac 1200  
 Asn Ala Gly Leu Ile Ala Glu Gly Gln Thr Glu Ile Thr Gly Val Asp  
                   385                  390                  395

aaa att cta aga ggc tac tct cat att gtt gaa aaa ctc aat gac cta 1248  
 Lys Ile Leu Arg Gly Tyr Ser His Ile Val Glu Lys Leu Asn Asp Leu  
                   400                  405                  410

ggc gca gat gtt tat atg caa gag ggg gaa gac tga 1284  
 Gly Ala Asp Val Tyr Met Gln Glu Gly Glu Asp  
                   415                  420                  425

<210> 6

<211> 425

<212> PRT

<213> *Alloiococcus otitidis*

<400> 6

Met Asn Ile Met Lys Lys Leu Ile Ile Asn Gly Gly Arg Thr Leu Lys  
                   1                  5                  10                  15

Gly Glu Val Thr Val Ser Gly Ala Lys Asn Ser Thr Val Ala Leu Ile  
                   20                  25                  30

Pro Ala Ser Ile Leu Ala Asp Ser Pro Val Ile Leu Glu Gly Val Pro  
                   35                  40                  45

Asp Ile Gln Asp Val His Ser Leu Leu Glu Ile Leu Asn Glu Met Asn  
                   50                  55                  60

Val Lys Thr Asp Phe Asp Gly Asn Thr Leu Thr Ile Asp Pro Arg Glu  
                   65                  70                  75                  80

Met Val Ser Ile Pro Met Pro Ser Gly Lys Ile Gln Ser Leu Arg Ala  
                   85                  90                  95

Ser Tyr Tyr Phe Met Gly Ala Leu Leu Ala Lys Phe Gly Lys Gly Val  
                   100                  105                  110

Val Gly Leu Pro Gly Gly Cys Phe Leu Gly Pro Arg Pro Ile Asp Gln  
                   115                  120                  125

His Leu Lys Gly Phe Arg Leu Leu Gly Ala Asp Val Asp Asn Glu Met  
                   130                  135                  140

Gly Ala Met Tyr Leu Lys Thr Ser Asp Ser Gly Leu Val Gly Ser Arg  
145 150 155 160

Ile Tyr Leu Asp Val Val Ser Ile Gly Ala Thr Ile Asn Ile Met Leu  
165 170 175

Ala Ala Val Arg Ala Gln Gly Arg Thr Val Ile Glu Asn Ala Ala Arg  
180 185 190

Glu Pro Glu Ile Ile Asp Val Ala Thr Leu Leu Asn Lys Met Gly Ala  
195 200 205

Lys Ile Arg Gly Ala Gly Thr Asp Met Ile Arg Ile Glu Gly Val Asp  
210 215 220

Gln Leu Thr Gly Cys Gln His Ser Ile Ile Pro Asp Arg Ile Glu Ala  
225 230 235 240

Gly Thr Tyr Leu Ala Ile Ala Ala Ala Gly Glu Asp Val Leu Val  
245 250 255

Asn Asn Val Ile Val Glu His Ile Asp Ser Leu Ile Ala Lys Leu Asp  
260 265 270

Glu Ile Gly Ile Asp Leu Asp Ile Gly Glu Asp Ser Ile Arg Val Lys  
275 280 285

Ala Pro Ser Lys Pro Leu Gln Pro Val Thr Ile Lys Thr Leu Pro Tyr  
290 295 300

Pro Gly Phe Ala Thr Asp Leu Gln Gln Pro Ile Thr Pro Leu Leu Leu  
305 310 315 320

Leu Ala Lys Gly Glu Ser Val Ile Thr Asp Thr Ile Tyr Pro Lys Arg  
325 330 335

Val Lys His Ile Pro Glu Leu Glu Arg Met Gly Ala Asn Ile Arg Val  
340 345 350

Glu Ser Asp Ile Ile Leu Ile Glu Gly Gly His Pro Leu Lys Gly Ala  
355 360 365

Glu Val Glu Ala Ser Asp Leu Arg Ala Gly Ala Cys Leu Ile Asn Ala  
 370 375 380

Gly Leu Ile Ala Glu Gly Gln Thr Glu Ile Thr Gly Val Asp Lys Ile  
 385 390 395 400

Leu Arg Gly Tyr Ser His Ile Val Glu Lys Leu Asn Asp Leu Gly Ala  
 405 410 415

Asp Val Tyr Met Gln Glu Gly Glu Asp  
 420 425

<210> 7

<211> 612

<212> DNA

<213> *Alloiococcus otitidis*

<220>

<221> CDS

<222> (4)..(612)

<223>

<400> 7

ctt ttg cat aga caa gac ttg aat cgt gaa agg aag tca gat gtg gaa 48  
 Met His Arg Gln Asp Leu Asn Arg Glu Arg Lys Ser Asp Val Glu  
 1 5 10 15

tta aaa gag ttt gat gga aag aaa aaa gaa gaa cta gcc atg att gat 96  
 Leu Lys Glu Phe Asp Gly Lys Lys Lys Glu Glu Leu Ala Met Ile Asp  
 20 25 30

gtg gcc aag gcc att tta gac cag gtc cat gac ttg atg cac ttc aac 144  
 Val Ala Lys Ala Ile Leu Asp Gln Val His Asp Leu Met His Phe Asn  
 35 40 45

gac ctc ttg agt gaa gtg tct gaa tat cta gac ttg tca gat gac gag 192  
 Asp Leu Leu Ser Glu Val Ser Glu Tyr Leu Asp Leu Ser Asp Asp Glu  
 50 55 60

atc gaa agc ggt atg ggc caa ttt tac acc gat tta aat att gac ggt 240  
 Ile Glu Ser Gly Met Gly Gln Phe Tyr Thr Asp Leu Asn Ile Asp Gly  
 65 70 75

cgc ttc atc tct tta ggc gac aac cat tgg ggc tta cgt gaa tgg tat 288  
 Arg Phe Ile Ser Leu Gly Asp Asn His Trp Gly Leu Arg Glu Trp Tyr  
 80 85 90 95

cca gtc gat tct atc gat gaa gag ttg acc cac gac aat gac ctg gag 336  
 Pro Val Asp Ser Ile Asp Glu Glu Leu Thr His Asp Asn Asp Leu Glu  
 100 105 110

aag gtc aca ccc aag cag gcg gaa gac ggc ttt gat gac tta gag cat 384

Lys Val Thr Pro Lys Gln Ala Glu Asp Gly Phe Asp Asp Leu Glu His  
 115 120 125

gtc gaa aaa gaa gtg atg gat gac gca aaa gaa gaa tta gat gac cag 432  
 Val Glu Lys Glu Val Met Asp Asp Ala Lys Glu Glu Leu Asp Asp Gln  
 130 135 140

gcc gtc aat gaa gat gaa gaa aat gtt gct cca gat gaa atc acc gac 480  
 Ala Val Asn Glu Asp Glu Glu Asn Val Ala Pro Asp Glu Ile Thr Asp  
 145 150 155

gat gga gat gaa gac aag ctg gat gaa tac tct agc gat atc gaa gac 528  
 Asp Gly Asp Glu Asp Lys Leu Asp Glu Tyr Ser Ser Asp Ile Glu Asp  
 160 165 170 175

ctc gaa gat gat cgt aag gct agc caa gac aag ctg tcc att gtt gac 576  
 Leu Glu Asp Asp Arg Lys Ala Ser Gln Asp Lys Leu Ser Ile Val Asp  
 180 185 190

gac gaa gat gtc tta aca aat gat gac gat gag taa 612  
 Asp Glu Asp Val Leu Thr Asn Asp Asp Asp Glu  
 195 200

<210> 8

<211> 202

<212> PRT

<213> Alloiococcus otitidis

<400> 8

Met His Arg Gln Asp Leu Asn Arg Glu Arg Lys Ser Asp Val Glu Leu  
 1 5 10 15

Lys Glu Phe Asp Gly Lys Lys Lys Glu Glu Leu Ala Met Ile Asp Val  
 20 25 30

Ala Lys Ala Ile Leu Asp Gln Val His Asp Leu Met His Phe Asn Asp  
 35 40 45

Leu Leu Ser Glu Val Ser Glu Tyr Leu Asp Leu Ser Asp Asp Glu Ile  
 50 55 60

Glu Ser Gly Met Gly Gln Phe Tyr Thr Asp Leu Asn Ile Asp Gly Arg  
 65 70 75 80

Phe Ile Ser Leu Gly Asp Asn His Trp Gly Leu Arg Glu Trp Tyr Pro  
 85 90 95

Val Asp Ser Ile Asp Glu Glu Leu Thr His Asp Asn Asp Leu Glu Lys  
 100 105 110

Val Thr Pro Lys Gln Ala Glu Asp Gly Phe Asp Asp Leu Glu His Val  
 115 120 125

Glu Lys Glu Val Met Asp Asp Ala Lys Glu Glu Leu Asp Asp Gln Ala  
 130 135 140

Val Asn Glu Asp Glu Glu Asn Val Ala Pro Asp Glu Ile Thr Asp Asp  
 145 150 155 160

Gly Asp Glu Asp Lys Leu Asp Glu Tyr Ser Ser Asp Ile Glu Asp Leu  
 165 170 175

Glu Asp Asp Arg Lys Ala Ser Gln Asp Lys Leu Ser Ile Val Asp Asp  
 180 185 190

Glu Asp Val Leu Thr Asn Asp Asp Asp Glu  
 195 200

<210> 9

<211> 942

<212> DNA

<213> Alloiococcus otitidis

<220>

<221> CDS

<222> (1) .. (942)

<223>

<400> 9

atg atc gaa att gaa aag cca gta att gaa aca gta gag atc agt gaa 48  
 Met Ile Glu Ile Glu Lys Pro Val Ile Glu Thr Val Glu Ile Ser Glu  
 1 5 10 15

gat ggc aaa ttc ggt aag ttt gtt gtt gaa cca ttg gaa cgt ggt tat 96  
 Asp Gly Lys Phe Gly Lys Phe Val Val Glu Pro Leu Glu Arg Gly Tyr  
 20 25 30

ggg act acc tta ggg aat tcc tta cgc cgc atc tta tta tca tca cta 144  
 Gly Thr Thr Leu Gly Asn Ser Leu Arg Arg Ile Leu Leu Ser Ser Leu  
 35 40 45

ccg ggt gct gcg gtc acc aat att caa att gat ggt gtt ttg cat gag 192  
 Pro Gly Ala Ala Val Thr Asn Ile Gln Ile Asp Gly Val Leu His Glu  
 50 55 60

ttt aca gct att gat ggt gtg gtt gaa gat gtg act tcc atc atc tta 240  
 Phe Thr Ala Ile Asp Gly Val Val Glu Asp Val Thr Ser Ile Ile Leu  
 65 70 75 80

aac ctg aaa aaa ctg gct tta aaa ctt cat act gaa gaa aca aaa aca 288



305

310

&lt;210&gt; 10

&lt;211&gt; 313

&lt;212&gt; PRT

<213> *Alloiococcus otitidis*

&lt;400&gt; 10

Met Ile Glu Ile Glu Lys Pro Val Ile Glu Thr Val Glu Ile Ser Glu  
 1 5 10 15

Asp Gly Lys Phe Gly Lys Phe Val Val Glu Pro Leu Glu Arg Gly Tyr  
 20 25 30

Gly Thr Thr Leu Gly Asn Ser Leu Arg Arg Ile Leu Leu Ser Ser Leu  
 35 40 45

Pro Gly Ala Ala Val Thr Asn Ile Gln Ile Asp Gly Val Leu His Glu  
 50 55 60

Phe Thr Ala Ile Asp Gly Val Val Glu Asp Val Thr Ser Ile Ile Leu  
 65 70 75 80

Asn Leu Lys Lys Leu Ala Leu Lys Leu His Thr Glu Glu Thr Lys Thr  
 85 90 95

Ile Glu Leu Asp Ile Glu Gly Pro Ala Glu Val Thr Ala Ala Asp Ile  
 100 105 110

Ile Thr Asp Ser Asp Val Glu Ile Met Asn Pro Asp Leu Tyr Leu Cys  
 115 120 125

Thr Val Ser Glu Gly Gly His Leu His Ile Arg Met Glu Ala Glu Thr  
 130 135 140

Gly Arg Gly Tyr Val Asn Ala Glu His Asn Lys His Asp Asp Met Pro  
 145 150 155 160

Ile Gly Val Leu Pro Ile Asp Ser Ile Tyr Thr Pro Ile Ser Arg Val  
 165 170 175

Asn Tyr Thr Val Glu Asp Thr Arg Val Gly Glu Arg Glu Gln Tyr Asp  
 180 185 190



Gly Leu Ser Leu Ala Ala Lys Ile Met Asn Glu His Leu Asn Ile Phe  
210 215 220

Ile Asn Leu Thr Glu Gln Ala Arg Glu Ala Asp Ile Met Val Glu Lys  
225 230 235 240

Glu Glu Asp Gln Lys Glu Lys Met Leu Glu Met Thr Ile Glu Glu Leu  
245 250 255

Asp Leu Ser Val Arg Ser Tyr Asn Cys Leu Lys Arg Ala Gly Ile Asn  
260 265 270

Thr Val Gln Glu Leu Thr Asp Lys Thr Glu Pro Glu Met Met Lys Val  
275 280 285

Arg Asn Leu Gly Arg Lys Ser Leu Glu Glu Val Lys Asn Lys Leu Asp  
290 295 300

Asp Leu Asp Leu Ser Leu Lys Glu Glu  
305 310

```
<210> 11
<211> 3681
<212> DNA
<213> Alloiococcus otitidis
```

```
<220>
<221> CDS
<222> (22) .. (3681)
<223>
```

```
<400> 11
aataaaggga ggtttgcccc c ttg gta gat gta aat aat ttt gaa agt att      51
      Met Val Asp Val Asn Asn Phe Glu Ser Ile
            1             5             10
```

caa att gga ctg gct tca cca gag aaa atc cgt tca tgg tct cat ggt 99  
Gln Ile Gly Leu Ala Ser Pro Glu Lys Ile Arg Ser Trp Ser His Gly  
15 20 25

gaa gtg aag aaa cct gaa acc att aac tac cgg aca tta aaa cct gaa 147  
Glu Val Lys Lys Pro Glu Thr Ile Asn Tyr Arg Thr Leu Lys Pro Glu  
30 35 40

aaa gac ggt ttg ttc tgc gaa cgc att ttt ggc cca acc aag gac tat 195  
Lys Asp Gly Leu Phe Cys Glu Arg Ile Phe Gly Pro Thr Lys Asp Tyr

45	50	55	
gaa tgt gct tgc gga aaa tat aaa cga gtc cac tat aaa ggg ata gtt			243
Glu Cys Ala Cys Gly Lys Tyr Lys Arg Val His Tyr Lys Gly Ile Val			
60	65	70	
tgt gac cgt tgc ggt gtt gaa gtc acc aag tgc agt gtc aga cga gaa			291
Cys Asp Arg Cys Gly Val Glu Val Thr Lys Ser Ser Val Arg Arg Glu			
75	80	85	90
cgc atg ggc cac ttg gaa tta gca gct cct gtc acc cac att tgg tac			339
Arg Met Gly His Leu Glu Leu Ala Ala Pro Val Thr His Ile Trp Tyr			
95	100	105	
ttc aag ggt att cca agt cgg atg ggc ctt atc tta gat atg agc cca			387
Phe Lys Gly Ile Pro Ser Arg Met Gly Leu Ile Leu Asp Met Ser Pro			
110	115	120	
aga tcc ttg gaa gaa att atc tat ttt gcc tct tat gtt gtt att gac			435
Arg Ser Leu Glu Glu Ile Ile Tyr Phe Ala Ser Tyr Val Val Ile Asp			
125	130	135	
ggt ggg gat acc cgc ctt gaa cgc aaa cag ctc tta act gaa cgt gaa			483
Gly Gly Asp Thr Pro Leu Glu Arg Lys Gln Leu Leu Thr Glu Arg Glu			
140	145	150	
tac cgg gaa aac aaa agc aag tac ggc aat gaa ttc caa gct gaa att			531
Tyr Arg Glu Asn Lys Ser Lys Tyr Gly Asn Glu Phe Gln Ala Glu Ile			
155	160	165	170
gga gct gaa gct gtt cgg acc ttg cta aaa aat gtc gat ttg gaa caa			579
Gly Ala Glu Ala Val Arg Thr Leu Leu Lys Asn Val Asp Leu Glu Gln			
175	180	185	
gaa gtt gct gac ctc aaa gaa atc tta gaa act gca act ggc caa aaa			627
Glu Val Ala Asp Leu Lys Glu Ile Leu Glu Thr Ala Thr Gly Gln Lys			
190	195	200	
cgg acc cgg gct att cgt cgt tta gac att att gac tcc ttc aag tct			675
Arg Thr Arg Ala Ile Arg Arg Leu Asp Ile Ile Asp Ser Phe Lys Ser			
205	210	215	
tcc aac aac aaa cgc gaa tgg atg gtc ttg gat gct att cca att atc			723
Ser Asn Asn Lys Pro Glu Trp Met Val Leu Asp Ala Ile Pro Ile Ile			
220	225	230	
cca cct gaa ctc cgc cca atg gta caa cta gaa ggt ggc cgg ttt gca			771
Pro Pro Glu Leu Arg Pro Met Val Gln Leu Glu Gly Gly Arg Phe Ala			
235	240	245	250
acc agc gac ttg aac gac ttg tac cgc cgg gtg att aac cgg aac aac			819
Thr Ser Asp Leu Asn Asp Leu Tyr Arg Arg Val Ile Asn Arg Asn Asn			
255	260	265	
cgg ttg aaa cgc ttg ctt gac ttg aat gcc ccc cac att atc gtc caa			867
Arg Leu Lys Arg Leu Leu Asp Leu Asn Ala Pro His Ile Ile Val Gln			
270	275	280	

aat gaa aaa cgg atg ctg caa gaa gct gtt gac gcc ttg att gac aat	915
Asn Glu Lys Arg Met Leu Gln Glu Ala Val Asp Ala Leu Ile Asp Asn	
285 290 295	
ggg cgt cgc ggt cgg gca gtc aac ggt cct ggt aac cgt ccg ctt aaa	963
Gly Arg Arg Gly Arg Ala Val Asn Gly Pro Gly Asn Arg Pro Leu Lys	
300 305 310	
tct ctt tct cac atg ttg aaa ggg aaa caa ggg cgc ttc cgt cag aac	1011
Ser Leu Ser His Met Leu Lys Gly Lys Gln Gly Arg Phe Arg Gln Asn	
315 320 325 330	
cta cta ggg aaa cgg gtt gac tac tct ggc cgg tct gtc att gtt gtt	1059
Leu Leu Gly Lys Arg Val Asp Tyr Ser Gly Arg Ser Val Ile Val Val	
335 340 345	
ggg cca acc ctt aaa atg tac caa tgt ggt cta ccg aaa gaa atg gcc	1107
Gly Pro Thr Leu Lys Met Tyr Gln Cys Gly Leu Pro Lys Glu Met Ala	
350 355 360	
atc gaa ctc ttc aaa cct ttt gtc atg cgg gag cta gtt gag cga gat	1155
Ile Glu Leu Phe Lys Pro Phe Val Met Arg Glu Leu Val Glu Arg Asp	
365 370 375	
att gca aat aac att aaa aat gcc aaa cga aaa gtg gaa cgg atg gaa	1203
Ile Ala Asn Asn Ile Lys Asn Ala Lys Arg Lys Val Glu Arg Met Glu	
380 385 390	
gat gat gtc tgg cct gtt tta gaa gat gtc att aaa gaa cac cct gtc	1251
Asp Asp Val Trp Pro Val Leu Glu Asp Val Ile Lys Glu His Pro Val	
395 400 405 410	
ctc tta aac cgg gcc cct acc ctt cac cgg cta ggg atc caa gcc ttt	1299
Leu Leu Asn Arg Ala Pro Thr Leu His Arg Leu Gly Ile Gln Ala Phe	
415 420 425	
gaa cct gtc ctt gtc aat ggg aag gct att cgc tta cac cca ctc gct	1347
Glu Pro Val Leu Val Asn Gly Lys Ala Ile Arg Leu His Pro Leu Ala	
430 435 440	
tgt gaa gcc tac aat gct gac ttt gac gga gac caa atg gct gtc cac	1395
Cys Glu Ala Tyr Asn Ala Asp Phe Asp Gly Asp Gln Met Ala Val His	
445 450 455	
gta ccc ctc agt gat gaa gcc cag gca gaa gcc cgc atc tta atg ctg	1443
Val Pro Leu Ser Asp Glu Ala Gln Ala Glu Ala Arg Ile Leu Met Leu	
460 465 470	
ggg gcc caa aat atc tta aac cct aaa gat ggt caa cca gtc gtt acc	1491
Gly Ala Gln Asn Ile Leu Asn Pro Lys Asp Gly Gln Pro Val Val Thr	
475 480 485 490	
cct tcc caa gac atg gtc cta ggg aac tac tac cta acc atg gaa gaa	1539
Pro Ser Gln Asp Met Val Leu Gly Asn Tyr Tyr Leu Thr Met Glu Glu	
495 500 505	

gaa ggt aaa att ggt gaa gga act gtc ttc tcc agt gct tct gag gct Glu Gly Lys Ile Gly Glu Gly Thr Val Phe Ser Ser Ala Ser Glu Ala 510 515 520	1587
atc caa gcc tac caa aca ggc tat gtc cac ctc cac acc cgg gtt gcg Ile Gln Ala Tyr Gln Thr Gly Tyr Val His Leu His Thr Arg Val Ala 525 530 535	1635
atc cgt gcg gtg gac tta ccg gac aaa cct ttt act gac tgg cag aaa Ile Arg Ala Val Asp Leu Pro Asp Lys Pro Phe Thr Asp Trp Gln Lys 540 545 550	1683
gac aag tac ttg att acc aca gtc ggt aag att atc ttt aat gaa att Asp Lys Tyr Leu Ile Thr Thr Val Gly Lys Ile Ile Phe Asn Glu Ile 555 560 565 570	1731
atg cca gca gaa ttt cca ttc ttg aac gaa cca tct aag gtt aac ctg Met Pro Ala Glu Phe Pro Phe Leu Asn Glu Pro Ser Lys Val Asn Leu 575 580 585	1779
gaa cag caa acc cca gac aag tac ttt gtc gac cgg ggc caa aac ttg Glu Gln Gln Thr Pro Asp Lys Tyr Phe Val Asp Arg Gly Gln Asn Leu 590 595 600	1827
aaa gac ctt att gcc gac cgt cct tta gtt cag cct ttc aaa aaa caa Lys Asp Leu Ile Ala Asp Arg Pro Leu Val Gln Pro Phe Lys Lys Gln 605 610 615	1875
gac ctg tcc aac att atc gcc gaa gtc ttt aat aac ttc caa gtg acc Asp Leu Ser Asn Ile Ile Ala Glu Val Phe Asn Asn Phe Gln Val Thr 620 625 630	1923
gaa acc tct aaa atg ttg gac cgc atg aag aac ttg ggc tac aag tac Glu Thr Ser Lys Met Leu Asp Arg Met Lys Asn Leu Gly Tyr Lys Tyr 635 640 645 650	1971
tct acc cgg tct ggt att act gtt ggg att gct gac gtt tca gtc cta Ser Thr Arg Ser Gly Ile Thr Val Gly Ile Ala Asp Val Ser Val Leu 655 660 665	2019
gaa gct aaa cca gaa atc ctg aaa gaa gcc cac gcc aag gtt gat aaa Glu Ala Lys Pro Glu Ile Leu Lys Glu Ala His Ala Lys Val Asp Lys 670 675 680	2067
atc aat gcc acc cac cgc cgc ggt tta att act gaa gaa gag cgt tac Ile Asn Ala Thr His Arg Arg Gly Leu Ile Thr Glu Glu Glu Arg Tyr 685 690 695	2115
gac aac gtt atc gat gtc tgg caa aag gct aag gat gaa att caa gat Asp Asn Val Ile Asp Val Trp Gln Lys Ala Lys Asp Glu Ile Gln Asp 700 705 710	2163
gcc ttg atg gat tcc ctt gac cca aga aat aac atc ttt atg atg tca Ala Leu Met Asp Ser Leu Asp Pro Arg Asn Asn Ile Phe Met Met Ser 715 720 725 730	2211
gac tct ggt gcc cgt ggg aat att tcc aac ttc acc caa cta gcc ggt	2259

Asp Ser Gly Ala Arg Gly Asn Ile Ser Asn Phe Thr Gln Leu Ala Gly	
735 740 745	
atg cgt ggt ttg atg gca gca cca agt ggt gag atc atg gaa ttg ccg	2307
Met Arg Gly Leu Met Ala Ala Pro Ser Gly Glu Ile Met Glu Leu Pro	
750 755 760	
atc acg tct aac ttc cgt gaa ggc ctg tct gtc tta gag atg ttt att	2355
Ile Thr Ser Asn Phe Arg Glu Gly Leu Ser Val Leu Glu Met Phe Ile	
765 770 775	
tcc acc cac ggt gcc cgt aaa ggc atg acc gat acc gcc ctt aaa act	2403
Ser Thr His Gly Ala Arg Lys Gly Met Thr Asp Thr Ala Leu Lys Thr	
780 785 790	
gcc gac tct ggt tac ttg acc aga cgt ttg gtt gat gtt gcc caa gac	2451
Ala Asp Ser Gly Tyr Leu Thr Arg Arg Leu Val Asp Val Ala Gln Asp	
795 800 805 810	
gtc atc atc cga gaa gaa gac tgt ggc act aaa cgt ggc ctt aaa gtt	2499
Val Ile Ile Arg Glu Glu Asp Cys Gly Thr Lys Arg Gly Leu Lys Val	
815 820 825	
tct gcc atc caa gta gga aat gaa cag att gaa agc ttg tct gac cgt	2547
Ser Ala Ile Gln Val Gly Asn Glu Gln Ile Glu Ser Leu Ser Asp Arg	
830 835 840	
atc ttg ggt cgt tat gcc caa gaa acc gtc acc cac ccc gaa act ggt	2595
Ile Leu Gly Arg Tyr Ala Gln Glu Thr Val Thr His Pro Glu Thr Gly	
845 850 855	
gaa gtc att gtt cac aag gat gaa ttg att gat gaa ggc aaa acc cga	2643
Glu Val Ile Val His Lys Asp Glu Leu Ile Asp Glu Gly Lys Thr Arg	
860 865 870	
aaa att gtc gat gcc ggt att gaa gaa gtt act atc cgg tct gcc ttc	2691
Lys Ile Val Asp Ala Gly Ile Glu Glu Val Thr Ile Arg Ser Ala Phe	
875 880 885 890	
tgc tgc aac acc aac cac ggt gtc tgc aag cac tgc tat ggc cgt aac	2739
Cys Cys Asn Thr Asn His Gly Val Cys Lys His Cys Tyr Gly Arg Asn	
895 900 905	
ttg gca act ggc cgg gaa gtt gaa gtt ggt gaa gca gtt gga act atc	2787
Leu Ala Thr Gly Arg Glu Val Glu Val Gly Glu Ala Val Gly Thr Ile	
910 915 920	
gct gcc caa tcc att ggg gaa ccc ggt acc caa ttg acc atg cgg acc	2835
Ala Ala Gln Ser Ile Gly Glu Pro Gly Thr Gln Leu Thr Met Arg Thr	
925 930 935	
ttc cac act ggt ggg gtc gct ggg gac gac atc acc caa ggt cta cca	2883
Phe His Thr Gly Gly Val Ala Gly Asp Asp Ile Thr Gln Gly Leu Pro	
940 945 950	
cgg gtt caa gaa atc ttt gaa gcc cgc cat ccg aaa ggg caa gcc acc	2931
Arg Val Gln Glu Ile Phe Glu Ala Arg His Pro Lys Gly Gln Ala Thr	

955	960	965	970	
att aca gaa gtg aat ggt caa atc caa gag atc gtt gaa gac cct gaa				2979
Ile Thr Glu Val Asn Gly Gln Ile Gln Glu Ile Val Glu Asp Pro Glu				
	975	980	985	
gaa cgc act aag acc gtc act gtt aag ggg aat gtt gac caa cgt gac				3027
Glu Arg Thr Lys Thr Val Thr Val Lys Gly Asn Val Asp Gln Arg Asp				
	990	995	1000	
tac tcc ttg cca atc aat gcc cgg atg aag gtt gaa gtt ggg gat tat				3075
Tyr Ser Leu Pro Ile Asn Ala Arg Met Lys Val Glu Val Gly Asp Tyr				
	1005	1010	1015	
gtt gaa cga ggc gat gct cta aac gag ggg tct att gat ccg aaa gag				3123
Val Glu Arg Gly Asp Ala Leu Asn Glu Gly Ser Ile Asp Pro Lys Glu				
	1020	1025	1030	
tta ctc gcg gtg agt gat atg atg aaa ttg cag aaa tac ctc ttg caa				3171
Leu Leu Ala Val Ser Asp Met Met Lys Leu Gln Lys Tyr Leu Leu Gln				
	1035	1040	1045	1050
gaa gtc caa tac gct tac cgg tct caa ggg gtc gaa att ggt gac aag				3219
Glu Val Gln Tyr Ala Tyr Arg Ser Gln Gly Val Glu Ile Gly Asp Lys				
	1055	1060	1065	
cac gtg gag gtt atg gtg cga caa atg ctc cgt aaa gtc cgt gtc ttg				3267
His Val Glu Val Met Val Arg Gln Met Leu Arg Lys Val Arg Val Leu				
	1070	1075	1080	
caa cca ggg gac act gat atc ctg cct ggt acc atg att gac ctc cac				3315
Gln Pro Gly Asp Thr Asp Ile Leu Pro Gly Thr Met Ile Asp Leu His				
	1085	1090	1095	
gac ttc aag gaa cgc aac caa gaa acc ttg atg tcc ggt ggc caa ccc				3363
Asp Phe Lys Glu Arg Asn Gln Glu Thr Leu Met Ser Gly Gly Gln Pro				
	1100	1105	1110	
gca act gct aga ctg gtc cta ctg ggt att acc aag gcc tcc ctt gaa				3411
Ala Thr Ala Arg Leu Val Leu Leu Gly Ile Thr Lys Ala Ser Leu Glu				
	1115	1120	1125	1130
acc aac tct ttc ttg tct gca gct tcc ttc caa gaa acc acc cgg gtc				3459
Thr Asn Ser Phe Leu Ser Ala Ala Ser Phe Gln Glu Thr Thr Arg Val				
	1135	1140	1145	
ctc acc gat gca gct att cgc ggt aaa gtt gat gac ctg gtt ggc ttg				3507
Leu Thr Asp Ala Ala Ile Arg Gly Lys Val Asp Asp Leu Val Gly Leu				
	1150	1155	1160	
aaa gaa aat gtt att atc ggt aaa tcc atc cca gct ggt act ggt atg				3555
Lys Glu Asn Val Ile Ile Gly Lys Ser Ile Pro Ala Gly Thr Gly Met				
	1165	1170	1175	
aga gcc tac agt aat att gaa cct aaa aaa gtt ggt gtc gtt agc gaa				3603
Arg Ala Tyr Ser Asn Ile Glu Pro Lys Lys Val Gly Val Val Ser Glu				
	1180	1185	1190	

aat gtc tac agc atc aat gaa gaa gac caa gtc agt caa gaa gaa aac 3651  
 Asn Val Tyr Ser Ile Asn Glu Glu Asp Gln Val Ser Gln Glu Glu Asn  
 1195 1200 1205 1210

cga gaa act gaa gaa act agc gag aaa taa 3681  
 Arg Glu Thr Glu Glu Thr Ser Glu Lys  
 1215

<210> 12

<211> 1219

<212> PRT

<213> *Alloioioccocus otitidis*

<400> 12

Met Val Asp Val Asn Asn Phe Glu Ser Ile Gln Ile Gly Leu Ala Ser  
 1 5 10 15

Pro Glu Lys Ile Arg Ser Trp Ser His Gly Glu Val Lys Lys Pro Glu  
 20 25 30

Thr Ile Asn Tyr Arg Thr Leu Lys Pro Glu Lys Asp Gly Leu Phe Cys  
 35 40 45

Glu Arg Ile Phe Gly Pro Thr Lys Asp Tyr Glu Cys Ala Cys Gly Lys  
 50 55 60

Tyr Lys Arg Val His Tyr Lys Gly Ile Val Cys Asp Arg Cys Gly Val  
 65 70 75 80

Glu Val Thr Lys Ser Ser Val Arg Arg Glu Arg Met Gly His Leu Glu  
 85 90 95

Leu Ala Ala Pro Val Thr His Ile Trp Tyr Phe Lys Gly Ile Pro Ser  
 100 105 110

Arg Met Gly Leu Ile Leu Asp Met Ser Pro Arg Ser Leu Glu Glu Ile  
 115 120 125

Ile Tyr Phe Ala Ser Tyr Val Val Ile Asp Gly Gly Asp Thr Pro Leu  
 130 135 140

Glu Arg Lys Gln Leu Leu Thr Glu Arg Glu Tyr Arg Glu Asn Lys Ser  
 145 150 155 160

Lys Tyr Gly Asn Glu Phe Gln Ala Glu Ile Gly Ala Glu Ala Val Arg

165

170

175

Thr Leu Leu Lys Asn Val Asp Leu Glu Gln Glu Val Ala Asp Leu Lys  
 180 185 190

Glu Ile Leu Glu Thr Ala Thr Gly Gln Lys Arg Thr Arg Ala Ile Arg  
 195 200 205

Arg Leu Asp Ile Ile Asp Ser Phe Lys Ser Ser Asn Asn Lys Pro Glu  
 210 215 220

Trp Met Val Leu Asp Ala Ile Pro Ile Ile Pro Pro Glu Leu Arg Pro  
 225 230 235 240

Met Val Gln Leu Glu Gly Gly Arg Phe Ala Thr Ser Asp Leu Asn Asp  
 245 250 255

Leu Tyr Arg Arg Val Ile Asn Arg Asn Asn Arg Leu Lys Arg Leu Leu  
 260 265 270

Asp Leu Asn Ala Pro His Ile Ile Val Gln Asn Glu Lys Arg Met Leu  
 275 280 285

Gln Glu Ala Val Asp Ala Leu Ile Asp Asn Gly Arg Arg Gly Arg Ala  
 290 295 300

Val Asn Gly Pro Gly Asn Arg Pro Leu Lys Ser Leu Ser His Met Leu  
 305 310 315 320

Lys Gly Lys Gln Gly Arg Phe Arg Gln Asn Leu Leu Gly Lys Arg Val  
 325 330 335

Asp Tyr Ser Gly Arg Ser Val Ile Val Val Gly Pro Thr Leu Lys Met  
 340 345 350

Tyr Gln Cys Gly Leu Pro Lys Glu Met Ala Ile Glu Leu Phe Lys Pro  
 355 360 365

Phe Val Met Arg Glu Leu Val Glu Arg Asp Ile Ala Asn Asn Ile Lys  
 370 375 380

Asn Ala Lys Arg Lys Val Glu Arg Met Glu Asp Asp Val Trp Pro Val  
 385 390 395 400



Leu Glu Asp Val Ile Lys Glu His Pro Val Leu Leu Asn Arg Ala Pro  
405 410 415

Thr Leu His Arg Leu Gly Ile Gln Ala Phe Glu Pro Val Leu Val Asn  
420 425 430

Gly Lys Ala Ile Arg Leu His Pro Leu Ala Cys Glu Ala Tyr Asn Ala  
435 440 445

Asp Phe Asp Gly Asp Gln Met Ala Val His Val Pro Leu Ser Asp Glu  
450 455 460

Ala Gln Ala Glu Ala Arg Ile Leu Met Leu Gly Ala Gln Asn Ile Leu  
465 470 475 480

Asn Pro Lys Asp Gly Gln Pro Val Val Thr Pro Ser Gln Asp Met Val  
485 490 495

Leu Gly Asn Tyr Tyr Leu Thr Met Glu Glu Glu Gly Lys Ile Gly Glu  
500 505 510

Gly Thr Val Phe Ser Ser Ala Ser Glu Ala Ile Gln Ala Tyr Gln Thr  
515 520 525

Gly Tyr Val His Leu His Thr Arg Val Ala Ile Arg Ala Val Asp Leu  
530 535 540

Pro Asp Lys Pro Phe Thr Asp Trp Gln Lys Asp Lys Tyr Leu Ile Thr  
545 550 555 560

Thr Val Gly Lys Ile Ile Phe Asn Glu Ile Met Pro Ala Glu Phe Pro  
565 570 575

Phe Leu Asn Glu Pro Ser Lys Val Asn Leu Glu Gln Gln Thr Pro Asp  
580 585 590

Lys Tyr Phe Val Asp Arg Gly Gln Asn Leu Lys Asp Leu Ile Ala Asp  
595 600 605

Arg Pro Leu Val Gln Pro Phe Lys Lys Gln Asp Leu Ser Asn Ile Ile  
610 615 620

Ala Glu Val Phe Asn Asn Phe Gln Val Thr Glu Thr Ser Lys Met Leu  
625 630 635 640

Asp Arg Met Lys Asn Leu Gly Tyr Lys Tyr Ser Thr Arg Ser Gly Ile  
645 650 655

Thr Val Gly Ile Ala Asp Val Ser Val Leu Glu Ala Lys Pro Glu Ile  
660 665 670

Leu Lys Glu Ala His Ala Lys Val Asp Lys Ile Asn Ala Thr His Arg  
675 680 685

Arg Gly Leu Ile Thr Glu Glu Glu Arg Tyr Asp Asn Val Ile Asp Val  
690 695 700

Trp Gln Lys Ala Lys Asp Glu Ile Gln Asp Ala Leu Met Asp Ser Leu  
705 710 715 720

Asp Pro Arg Asn Asn Ile Phe Met Met Ser Asp Ser Gly Ala Arg Gly  
725 730 735

Asn Ile Ser Asn Phe Thr Gln Leu Ala Gly Met Arg Gly Leu Met Ala  
740 745 750

Ala Pro Ser Gly Glu Ile Met Glu Leu Pro Ile Thr Ser Asn Phe Arg  
755 760 765

Glu Gly Leu Ser Val Leu Glu Met Phe Ile Ser Thr His Gly Ala Arg  
770 775 780

Lys Gly Met Thr Asp Thr Ala Leu Lys Thr Ala Asp Ser Gly Tyr Leu  
785 790 795 800

Thr Arg Arg Leu Val Asp Val Ala Gln Asp Val Ile Ile Arg Glu Glu  
805 810 815

Asp Cys Gly Thr Lys Arg Gly Leu Lys Val Ser Ala Ile Gln Val Gly  
820 825 830

Asn Glu Gln Ile Glu Ser Leu Ser Asp Arg Ile Leu Gly Arg Tyr Ala  
835 840 845

Gln Glu Thr Val Thr His Pro Glu Thr Gly Glu Val Ile Val His Lys  
850 855 860

Asp Glu Leu Ile Asp Glu Gly Lys Thr Arg Lys Ile Val Asp Ala Gly  
865 870 875 880

Ile Glu Glu Val Thr Ile Arg Ser Ala Phe Cys Cys Asn Thr Asn His  
885 890 895

Gly Val Cys Lys His Cys Tyr Gly Arg Asn Leu Ala Thr Gly Arg Glu  
900 905 910

Val Glu Val Gly Glu Ala Val Gly Thr Ile Ala Ala Gln Ser Ile Gly  
915 920 925

Glu Pro Gly Thr Gln Leu Thr Met Arg Thr Phe His Thr Gly Gly Val  
930 935 940

Ala Gly Asp Asp Ile Thr Gln Gly Leu Pro Arg Val Gln Glu Ile Phe  
945 950 955 960

Glu Ala Arg His Pro Lys Gly Gln Ala Thr Ile Thr Glu Val Asn Gly  
965 970 975

Gln Ile Gln Glu Ile Val Glu Asp Pro Glu Glu Arg Thr Lys Thr Val  
980 985 990

Thr Val Lys Gly Asn Val Asp Gln Arg Asp Tyr Ser Leu Pro Ile Asn  
995 1000 1005

Ala Arg Met Lys Val Glu Val Gly Asp Tyr Val Glu Arg Gly Asp Ala  
1010 1015 1020

Leu Asn Glu Gly Ser Ile Asp Pro Lys Glu Leu Leu Ala Val Ser Asp  
1025 1030 1035 1040

Met Met Lys Leu Gln Lys Tyr Leu Leu Gln Glu Val Gln Tyr Ala Tyr  
1045 1050 1055

Arg Ser Gln Gly Val Glu Ile Gly Asp Lys His Val Glu Val Met Val  
1060 1065 1070

Arg Gln Met Leu Arg Lys Val Arg Val Leu Gln Pro Gly Asp Thr Asp

1075

1080

1085

Ile Leu Pro Gly Thr Met Ile Asp Leu His Asp Phe Lys Glu Arg Asn  
 1090 1095 1100

Gln Glu Thr Leu Met Ser Gly Gly Gln Pro Ala Thr Ala Arg Leu Val  
 1105 1110 1115 1120

Leu Leu Gly Ile Thr Lys Ala Ser Leu Glu Thr Asn Ser Phe Leu Ser  
 1125 1130 1135

Ala Ala Ser Phe Gln Glu Thr Thr Arg Val Leu Thr Asp Ala Ala Ile  
 1140 1145 1150

Arg Gly Lys Val Asp Asp Leu Val Gly Leu Lys Glu Asn Val Ile Ile  
 1155 1160 1165

Gly Lys Ser Ile Pro Ala Gly Thr Gly Met Arg Ala Tyr Ser Asn Ile  
 1170 1175 1180

Glu Pro Lys Lys Val Gly Val Val Ser Glu Asn Val Tyr Ser Ile Asn  
 1185 1190 1195 1200

Glu Glu Asp Gln Val Ser Gln Glu Glu Asn Arg Glu Thr Glu Glu Thr  
 1205 1210 1215

Ser Glu Lys

<210> 13

<211> 3582

<212> DNA

<213> *Alloioococcus otitidis*

<220>

<221> CDS

<222> (4)..(3582)

<223>

<400> 13

gag gtg aac aag ttg gtc ggt aaa aaa gtt aat ttt ggt aaa cac cgt  
 Met Asn Lys Leu Val Gly Lys Lys Val Asn Phe Gly Lys His Arg  
 1 5 10 15

48

gtt cgt aga agt tac tca cga atc aac gaa gta ctc gag ctc ccg aat  
 Val Arg Arg Ser Tyr Ser Arg Ile Asn Glu Val Leu Glu Leu Pro Asn  
 20 25 30

96

tta att gaa atc cag act gat tca tat gat tgg ttt tta gat gaa ggc Leu Ile Glu Ile Gln Thr Asp Ser Tyr Asp Trp Phe Leu Asp Glu Gly 35 40 45	144
ttg aag gaa atg ttt agt gat att tcc cca atc gat gat ttt tca ggc Leu Lys Glu Met Phe Ser Asp Ile Ser Pro Ile Asp Asp Phe Ser Gly 50 55 60	192
aat ttg tcc cta gaa ttt gtt gac tat aaa ttt tac gaa agc aag tat Asn Leu Ser Leu Glu Phe Val Asp Tyr Lys Phe Tyr Glu Ser Lys Tyr 65 70 75	240
act gtt gaa gaa gct aga gag cat gat gcg aac tat tct gcc ccc ctc Thr Val Glu Glu Ala Arg Glu His Asp Ala Asn Tyr Ser Ala Pro Leu 80 85 90 95	288
tac gtg aag tta cgt ttg atc aac aag gaa act ggt gaa gtc aag gaa Tyr Val Lys Leu Arg Leu Ile Asn Lys Glu Thr Gly Glu Val Lys Glu 100 105 110	336
caa gaa gtc ttc ttc ggt gac ttt ccg tta atg aca gaa caa ggg acc Gln Glu Val Phe Gly Asp Phe Pro Leu Met Thr Glu Gln Gly Thr 115 120 125	384
ttt atc atc aac ggg gct gag cgg gtg att gtt tcc caa ctt gtc cgg Phe Ile Ile Asn Gly Ala Glu Arg Val Ile Val Ser Gln Leu Val Arg 130 135 140	432
tcg cct ggg gtt tat tac agt cca aaa gtt gag aaa aac ggc cgg gaa Ser Pro Gly Val Tyr Tyr Ser Pro Lys Val Glu Lys Asn Gly Arg Glu 145 150 155	480
ggt ttt tca acc gtt ctt atc cct aac cgg ggt gct tgg ctt gaa tac Gly Phe Ser Thr Val Leu Ile Pro Asn Arg Gly Ala Trp Leu Glu Tyr 160 165 170 175	528
gaa aca gat acc aaa ggc atc tcc aat gtt cga att gac cga acc cgt Glu Thr Asp Thr Lys Gly Ile Ser Asn Val Arg Ile Asp Arg Thr Arg 180 185 190	576
aaa att ccg atc act gtc ttg tta aga gct cta ggg att ggg tca gat Lys Ile Pro Ile Thr Val Leu Leu Arg Ala Leu Gly Ile Gly Ser Asp 195 200 205	624
gat gaa att att gac ctg atc ggc ttg aat gac agc ttg gaa gcc acc Asp Glu Ile Ile Asp Leu Ile Gly Leu Asn Asp Ser Leu Glu Ala Thr 210 215 220	672
ttg gaa aag gat gtc cac aag tct act tca gat tcc cgc gta gaa gaa Leu Glu Lys Asp Val His Lys Ser Thr Ser Asp Ser Arg Val Glu Glu 225 230 235	720
gcc ttg aaa gac ttg tat gaa cgc ttg cgt cca ggt gaa ccc aaa aca Ala Leu Lys Asp Leu Tyr Glu Arg Leu Arg Pro Gly Glu Pro Lys Thr 240 245 250 255	768
gct gaa tcc tct cgt aac ttg atc aat acc cgg ttc ttt gac cac aag	816

Ala Glu Ser Ser Arg Asn Leu Ile Asn Thr Arg Phe Phe Asp His Lys	
260 265 270	
cgt tac gac cta gcc tat gtt ggt cgc tac aag atg aac aaa aaa cta	864
Arg Tyr Asp Leu Ala Tyr Val Gly Arg Tyr Lys Met Asn Lys Lys Leu	
275 280 285	
gac ctt aaa acc cgc ttg atg ggg act gtc ctt gcc gaa aac ctg gtt	912
Asp Leu Lys Thr Arg Leu Met Gly Thr Val Leu Ala Glu Asn Leu Val	
290 295 300	
gat cct gaa gct ggc gag gtc tta gct gaa gaa ggt agt gaa gtg acc	960
Asp Pro Glu Ala Gly Glu Val Leu Ala Glu Glu Gly Ser Glu Val Thr	
305 310 315	
cgg tct gtg atg gac aag ctt ggc cct tac ctt gac ggt gac atg aac	1008
Arg Ser Val Met Asp Lys Leu Gly Pro Tyr Leu Asp Gly Asp Met Asn	
320 325 330 335	
caa gta acc att aac ccc tca gaa gaa gcg gtt atc cct gaa ccc att	1056
Gln Val Thr Ile Asn Pro Ser Glu Glu Ala Val Ile Pro Glu Pro Ile	
340 345 350	
gac cta caa att gtc aaa gtc tac tcc aaa gaa gat cca gac cgg atc	1104
Asp Leu Gln Ile Val Lys Val Tyr Ser Lys Glu Asp Pro Asp Arg Ile	
355 360 365	
gtg aac atg atc ggc aac ggg cac cct gac aaa aag gcc aaa tgg att	1152
Val Asn Met Ile Gly Asn Gly His Pro Asp Lys Lys Ala Lys Trp Ile	
370 375 380	
acc cct gct gac atg ata gcg gct atg agt tac ttc ttt aac ctc caa	1200
Thr Pro Ala Asp Met Ile Ala Ala Met Ser Tyr Phe Phe Asn Leu Gln	
385 390 395	
gaa ggc att ggc gat gtt gac gat atc gac cac ttg ggt aac cgt cgg	1248
Glu Gly Ile Gly Asp Val Asp Asp Ile Asp His Leu Gly Asn Arg Arg	
400 405 410 415	
atc cgg tca gtc gga gag ctt ttg caa aac caa ttc cga att ggg ctc	1296
Ile Arg Ser Val Gly Glu Leu Leu Gln Asn Gln Phe Arg Ile Gly Leu	
420 425 430	
tct cgg atg gag cgg gtg gtc cgc gaa cga atg tcc atc caa gac att	1344
Ser Arg Met Glu Arg Val Val Arg Glu Arg Met Ser Ile Gln Asp Ile	
435 440 445	
tct agc acc aca ccc caa caa tta att aac atc cgt ccc gtt gta gct	1392
Ser Ser Thr Thr Pro Gln Gln Leu Ile Asn Ile Arg Pro Val Val Ala	
450 455 460	
tct ctg aaa gaa ttt ttc ggg tct tcc caa ctc tcc caa ttc atg gac	1440
Ser Leu Lys Glu Phe Phe Gly Ser Ser Gln Leu Ser Gln Phe Met Asp	
465 470 475	
caa acc aac ccc ttg ggt gag tta acc cac aaa cgt cgc ttg tct gcc	1488
Gln Thr Asn Pro Leu Gly Glu Leu Thr His Lys Arg Arg Leu Ser Ala	

480	485	490	495	
ctt gga cca gga ggc ttg act agg gac cgg gct ggt tat gaa gtc cga				1536
Leu Gly Pro Gly Gly Leu Thr Arg Asp Arg Ala Gly Tyr Glu Val Arg				
500		505	510	
gac gtc cac tat tcc cac tac ggc cgg atg tgc ccg atc gaa aca cct				1584
Asp Val His Tyr Ser His Tyr Gly Arg Met Cys Pro Ile Glu Thr Pro				
515		520	525	
gaa ggc cca aac att ggt ctg att aac agt ttg tct acc tat gct aag				1632
Glu Gly Pro Asn Ile Gly Leu Ile Asn Ser Leu Ser Thr Tyr Ala Lys				
530		535	540	
atc aat aaa ttt ggt ttt att gaa aca cct tac cgc cgg gtg gac cgg				1680
Ile Asn Lys Phe Gly Phe Ile Glu Thr Pro Tyr Arg Arg Val Asp Arg				
545		550	555	
gaa act ggc cag gta acg gat aaa att gac tac ttg act gct gac gaa				1728
Glu Thr Gly Gln Val Thr Asp Lys Ile Asp Tyr Leu Thr Ala Asp Glu				
560		565	570	575
gaa gat ctt tac gtt gta gcc caa gcc aat gct gaa tta gat gaa gat				1776
Glu Asp Leu Tyr Val Val Ala Gln Ala Asn Ala Glu Leu Asp Glu Asp				
580		585	590	
gga cat ttt gct aat gat gtc gtc cta gcc cga aga cgg gat gtc aac				1824
Gly His Phe Ala Asn Asp Val Val Leu Ala Arg Arg Arg Asp Val Asn				
595		600	605	
gaa gag gtt gac gct tcc gaa gtt gac tat atg gac gtg tca cca aaa				1872
Glu Glu Val Asp Ala Ser Glu Val Asp Tyr Met Asp Val Ser Pro Lys				
610		615	620	
caa gtt gtt tct gtg gcc aca gct tcc att cct ttc tta gaa aac gac				1920
Gln Val Val Ser Val Ala Thr Ala Ser Ile Pro Phe Leu Glu Asn Asp				
625		630	635	
gac tcc aac cgg gct cta atg ggg gct aac atg caa cgg caa gct gtt				1968
Asp Ser Asn Arg Ala Leu Met Gly Ala Asn Met Gln Arg Gln Ala Val				
640		645	650	655
cct ctt atg caa cca gag tcc cca cta gta gga act gga atc gaa cac				2016
Pro Leu Met Gln Pro Glu Ser Pro Leu Val Gly Thr Gly Ile Glu His				
660		665	670	
att gca gcc cgt gac tct gga gct gcc gtt att gcc aag gct gac ggg				2064
Ile Ala Ala Arg Asp Ser Gly Ala Ala Val Ile Ala Lys Ala Asp Gly				
675		680	685	
gtt gtg gag tat gtt gat gcc aag acg gtc aaa gtc cgt caa gcc gat				2112
Val Val Glu Tyr Val Asp Ala Lys Thr Val Lys Val Arg Gln Ala Asp				
690		695	700	
ggt acc ctc aac aac tac aag ctg gct aag tac aaa cgg tcc aac tcc				2160
Gly Thr Leu Asn Asn Tyr Lys Leu Ala Lys Tyr Lys Arg Ser Asn Ser				
705		710	715	

gga act tct tac aac caa aga cca att gta aaa act ggt gag gaa gtt Gly Thr Ser Tyr Asn Gln Arg Pro Ile Val Lys Thr Gly Glu Glu Val 720 725 730 735	2208
gac aaa ggc gac atc cta gca gat ggt ccg tcc atg gaa aat ggt gaa Asp Lys Gly Asp Ile Leu Ala Asp Gly Pro Ser Met Glu Asn Gly Glu 740 745 750	2256
atg gcc ctt ggt aaa aac cca ttg att gcc ttt acc acc ttt gat ggc Met Ala Leu Gly Lys Asn Pro Leu Ile Ala Phe Thr Thr Phe Asp Gly 755 760 765	2304
tac aac ttc gag gat gcc gtc att atg agt gaa cgt ttg gtc aaa gat Tyr Asn Phe Glu Asp Ala Val Ile Met Ser Glu Arg Leu Val Lys Asp 770 775 780	2352
gac gtt tat acc tcc atc cac att gaa gaa tat gaa tct gaa gcc cgc Asp Val Tyr Thr Ser Ile His Ile Glu Glu Tyr Glu Ser Glu Ala Arg 785 790 795	2400
gat acc aag tta ggg cca gaa gaa atc acc cgg gaa att cca aac gtc Asp Thr Lys Leu Gly Pro Glu Glu Ile Thr Arg Glu Ile Pro Asn Val 800 805 810 815	2448
ggg gaa agt gcc ctc aag aac ttg gat gaa aga ggc att atc cgg atc Gly Glu Ser Ala Leu Lys Asn Leu Asp Glu Arg Gly Ile Ile Arg Ile 820 825 830	2496
ggg gct gaa gtt cgt gac ggg gac atc cta gtt ggt aaa gtt aca ccc Gly Ala Glu Val Arg Asp Gly Asp Ile Leu Val Gly Lys Val Thr Pro 835 840 845	2544
aaa ggg gtt agt gaa cta tca gct gag gaa aaa ctc ctc cac gct atc Lys Gly Val Ser Glu Leu Ser Ala Glu Glu Lys Leu Leu His Ala Ile 850 855 860	2592
ttc ggc gaa aaa gcc cgg gaa gtt cgt gac acc tcc ctc cgt gtc cca Phe Gly Glu Lys Ala Arg Glu Val Arg Asp Thr Ser Leu Arg Val Pro 865 870 875	2640
cac ggt agt ggc gga att gtc cac gat gtc cag atc ttt acc cgg gaa His Gly Ser Gly Gly Ile Val His Asp Val Gln Ile Phe Thr Arg Glu 880 885 890 895	2688
gcc ggc gac gaa ctg tca cca ggc gtt aac tac ctt gtc cga gtt ttc Ala Gly Asp Glu Leu Ser Pro Gly Val Asn Tyr Leu Val Arg Val Phe 900 905 910	2736
att gcc caa aaa cgt aaa att gac gtt ggg gac aag atg gca ggt cga Ile Ala Gln Lys Arg Lys Ile Asp Val Gly Asp Lys Met Ala Gly Arg 915 920 925	2784
cac ggg aac aag ggt gtt gtt tcc ctt atc tta cca gaa gaa gac atg His Gly Asn Lys Gly Val Val Ser Leu Ile Leu Pro Glu Glu Asp Met 930 935 940	2832



ccg ttt atg cca gac gga acc cca att gac atc atg ctc aac cca ctt Pro Phe Met Pro Asp Gly Thr Pro Ile Asp Ile Met Leu Asn Pro Leu 945 950 955	2880
ggt gtc cct tcc cgg atg aat gtc ggc cag gtc atc gaa ctc cac atg Gly Val Pro Ser Arg Met Asn Val Gly Gln Val Ile Glu Leu His Met 960 965 970 975	2928
ggg atg gca gcc cga cag tta ggc gag cac att gct act cca gtc ttt Gly Met Ala Ala Arg Gln Leu Gly Glu His Ile Ala Thr Pro Val Phe 980 985 990	2976
gac ggg gcc aac gaa gaa gat gtt tgg gaa act atc aag gaa gcc ggt Asp Gly Ala Asn Glu Glu Asp Val Trp Glu Thr Ile Lys Glu Ala Gly 995 1000 1005	3024
atg gat gca gat gcc aaa acc gtc ttg tat gac ggc cgg act ggc gag Met Asp Ala Asp Ala Lys Thr Val Leu Tyr Asp Gly Arg Thr Gly Glu 1010 1015 1020	3072
cca ttt gac aac aag gtc tcc gtt ggg gtg atg tac ttt atc aaa cta Pro Phe Asp Asn Lys Val Ser Val Gly Val Met Tyr Phe Ile Lys Leu 1025 1030 1035	3120
gtc cac atg gtc gac gac aag ttg cac gcc aga tcc aca gga cca tac Val His Met Val Asp Asp Lys Leu His Ala Arg Ser Thr Gly Pro Tyr 1040 1045 1050 1055	3168
tcc ttg gtt acc caa caa ccc ctt ggt ggg aaa gca cag ttt ggt ggc Ser Leu Val Thr Gln Gln Pro Leu Gly Gly Lys Ala Gln Phe Gly Gly 1060 1065 1070	3216
caa cgc ttt ggt gag atg gaa gtc tgg gcc ttg gaa gct tat ggg gct Gln Arg Phe Gly Glu Met Glu Val Trp Ala Leu Glu Ala Tyr Gly Ala 1075 1080 1085	3264
tcc cgc acc ctc caa gaa atc ttg acc tac aag tca gat gac gtg att Ser Arg Thr Leu Gln Glu Ile Leu Thr Tyr Lys Ser Asp Asp Val Ile 1090 1095 1100	3312
gga agg gta gac acc tat gaa gcc att gtc aag ggc caa cgc att cca Gly Arg Val Asp Thr Tyr Glu Ala Ile Val Lys Gly Gln Arg Ile Pro 1105 1110 1115	3360
aaa cct ggt gta cct gaa tcc ttc cgt gtc ctc gtg aaa gaa ctc cag Lys Pro Gly Val Pro Glu Ser Phe Arg Val Leu Val Lys Glu Leu Gln 1120 1125 1130 1135	3408
tct ctg ggg ttg gac ctg aaa gtc ctc gac aag gaa caa aac gaa atc Ser Leu Gly Leu Asp Leu Lys Val Leu Asp Lys Glu Gln Asn Glu Ile 1140 1145 1150	3456
aat ctc aag gct gaa gat gac gag tcg gaa gac caa gtc gtt gat tcc Asn Leu Lys Ala Glu Asp Asp Glu Ser Glu Asp Gln Val Val Asp Ser 1155 1160 1165	3504
cta gaa gaa atg cgt aaa gag cag gaa gaa gaa cgc cgt aag gaa aaa	3552

Leu Glu Glu Met Arg Lys Glu Gln Glu Glu Glu Arg Arg Lys Glu Lys  
 1170 1175 1180

gaa aaa gaa gag cca agt act gag tcg taa  
 Glu Lys Glu Glu Pro Ser Thr Glu Ser  
 1185 1190

3582

<210> 14  
 <211> 1192  
 <212> PRT  
 <213> Alloiococcus otitidis

<400> 14  
 Met Asn Lys Leu Val Gly Lys Lys Val Asn Phe Gly Lys His Arg Val  
 1 5 10 15

Arg Arg Ser Tyr Ser Arg Ile Asn Glu Val Leu Glu Leu Pro Asn Leu  
 20 25 30

Ile Glu Ile Gln Thr Asp Ser Tyr Asp Trp Phe Leu Asp Glu Gly Leu  
 35 40 45

Lys Glu Met Phe Ser Asp Ile Ser Pro Ile Asp Asp Phe Ser Gly Asn  
 50 55 60

Leu Ser Leu Glu Phe Val Asp Tyr Lys Phe Tyr Glu Ser Lys Tyr Thr  
 65 70 75 80

Val Glu Glu Ala Arg Glu His Asp Ala Asn Tyr Ser Ala Pro Leu Tyr  
 85 90 95

Val Lys Leu Arg Leu Ile Asn Lys Glu Thr Gly Glu Val Lys Glu Gln  
 100 105 110

Glu Val Phe Phe Gly Asp Phe Pro Leu Met Thr Glu Gln Gly Thr Phe  
 115 120 125

Ile Ile Asn Gly Ala Glu Arg Val Ile Val Ser Gln Leu Val Arg Ser  
 130 135 140

Pro Gly Val Tyr Tyr Ser Pro Lys Val Glu Lys Asn Gly Arg Glu Gly  
 145 150 155 160

Phe Ser Thr Val Leu Ile Pro Asn Arg Gly Ala Trp Leu Glu Tyr Glu  
 165 170 175

Thr Asp Thr Lys Gly Ile Ser Asn Val Arg Ile Asp Arg Thr Arg Lys  
180 185 190

Ile Pro Ile Thr Val Leu Leu Arg Ala Leu Gly Ile Gly Ser Asp Asp  
195 200 205

Glu Ile Ile Asp Leu Ile Gly Leu Asn Asp Ser Leu Glu Ala Thr Leu  
210 215 220

Glu Lys Asp Val His Lys Ser Thr Ser Asp Ser Arg Val Glu Glu Ala  
225 230 235 240

Leu Lys Asp Leu Tyr Glu Arg Leu Arg Pro Gly Glu Pro Lys Thr Ala  
245 250 255

Glu Ser Ser Arg Asn Leu Ile Asn Thr Arg Phe Phe Asp His Lys Arg  
260 265 270

Tyr Asp Leu Ala Tyr Val Gly Arg Tyr Lys Met Asn Lys Lys Leu Asp  
275 280 285

Leu Lys Thr Arg Leu Met Gly Thr Val Leu Ala Glu Asn Leu Val Asp  
290 295 300

Pro Glu Ala Gly Glu Val Leu Ala Glu Glu Gly Ser Glu Val Thr Arg  
305 310 315 320

Ser Val Met Asp Lys Leu Gly Pro Tyr Leu Asp Gly Asp Met Asn Gln  
325 330 335

Val Thr Ile Asn Pro Ser Glu Glu Ala Val Ile Pro Glu Pro Ile Asp  
340 345 350

Leu Gln Ile Val Lys Val Tyr Ser Lys Glu Asp Pro Asp Arg Ile Val  
355 360 365

Asn Met Ile Gly Asn Gly His Pro Asp Lys Lys Ala Lys Trp Ile Thr  
370 375 380

Pro Ala Asp Met Ile Ala Ala Met Ser Tyr Phe Phe Asn Leu Gln Glu  
385 390 395 400

Gly Ile Gly Asp Val Asp Asp Ile Asp His Leu Gly Asn Arg Arg Ile  
405 410 415

Arg Ser Val Gly Glu Leu Leu Gln Asn Gln Phe Arg Ile Gly Leu Ser  
420 425 430

Arg Met Glu Arg Val Val Arg Glu Arg Met Ser Ile Gln Asp Ile Ser  
435 440 445

Ser Thr Thr Pro Gln Gln Leu Ile Asn Ile Arg Pro Val Val Ala Ser  
450 455 460

Leu Lys Glu Phe Phe Gly Ser Ser Gln Leu Ser Gln Phe Met Asp Gln  
465 470 475 480

Thr Asn Pro Leu Gly Glu Leu Thr His Lys Arg Arg Leu Ser Ala Leu  
485 490 495

Gly Pro Gly Gly Leu Thr Arg Asp Arg Ala Gly Tyr Glu Val Arg Asp  
500 505 510

Val His Tyr Ser His Tyr Gly Arg Met Cys Pro Ile Glu Thr Pro Glu  
515 520 525

Gly Pro Asn Ile Gly Leu Ile Asn Ser Leu Ser Thr Tyr Ala Lys Ile  
530 535 540

Asn Lys Phe Gly Phe Ile Glu Thr Pro Tyr Arg Arg Val Asp Arg Glu  
545 550 555 560

Thr Gly Gln Val Thr Asp Lys Ile Asp Tyr Leu Thr Ala Asp Glu Glu  
565 570 575

Asp Leu Tyr Val Val Ala Gln Ala Asn Ala Glu Leu Asp Glu Asp Gly  
580 585 590

His Phe Ala Asn Asp Val Val Leu Ala Arg Arg Arg Asp Val Asn Glu  
595 600 605

Glu Val Asp Ala Ser Glu Val Asp Tyr Met Asp Val Ser Pro Lys Gln  
610 615 620

Val Val Ser Val Ala Thr Ala Ser Ile Pro Phe Leu Glu Asn Asp Asp

625                      630                      635                      640

Ser Asn Arg Ala Leu Met Gly Ala Asn Met Gln Arg Gln Ala Val Pro  
645                                      650                                      655

Leu Met Gln Pro Glu Ser Pro Leu Val Gly Thr Gly Ile Glu His Ile  
660                                      665                                      670

Ala Ala Arg Asp Ser Gly Ala Ala Val Ile Ala Lys Ala Asp Gly Val  
675                                      680                                      685

Val Glu Tyr Val Asp Ala Lys Thr Val Lys Val Arg Gln Ala Asp Gly  
690                                      695                                      700

Thr Leu Asn Asn Tyr Lys Leu Ala Lys Tyr Lys Arg Ser Asn Ser Gly  
705                                      710                                      715                                      720

Thr Ser Tyr Asn Gln Arg Pro Ile Val Lys Thr Gly Glu Glu Val Asp  
725                                      730                                      735

Lys Gly Asp Ile Leu Ala Asp Gly Pro Ser Met Glu Asn Gly Glu Met  
740                                      745                                      750

Ala Leu Gly Lys Asn Pro Leu Ile Ala Phe Thr Thr Phe Asp Gly Tyr  
755                                      760                                      765

Asn Phe Glu Asp Ala Val Ile Met Ser Glu Arg Leu Val Lys Asp Asp  
770                                      775                                      780

Val Tyr Thr Ser Ile His Ile Glu Glu Tyr Glu Ser Glu Ala Arg Asp  
785                                      790                                      795                                      800

Thr Lys Leu Gly Pro Glu Glu Ile Thr Arg Glu Ile Pro Asn Val Gly  
805                                      810                                      815

Glu Ser Ala Leu Lys Asn Leu Asp Glu Arg Gly Ile Ile Arg Ile Gly  
820                                      825                                      830

Ala Glu Val Arg Asp Gly Asp Ile Leu Val Gly Lys Val Thr Pro Lys  
835                                      840                                      845

Gly Val Ser Glu Leu Ser Ala Glu Glu Lys Leu Leu His Ala Ile Phe  
850                                      855                                      860

Gly Glu Lys Ala Arg Glu Val Arg Asp Thr Ser Leu Arg Val Pro His  
865 870 875 880

Gly Ser Gly Gly Ile Val His Asp Val Gln Ile Phe Thr Arg Glu Ala  
885 890 895

Gly Asp Glu Leu Ser Pro Gly Val Asn Tyr Leu Val Arg Val Phe Ile  
900 905 910

Ala Gln Lys Arg Lys Ile Asp Val Gly Asp Lys Met Ala Gly Arg His  
915 920 925

Gly Asn Lys Gly Val Val Ser Leu Ile Leu Pro Glu Glu Asp Met Pro  
930 935 940

Phe Met Pro Asp Gly Thr Pro Ile Asp Ile Met Leu Asn Pro Leu Gly  
945 950 955 960

Val Pro Ser Arg Met Asn Val Gly Gln Val Ile Glu Leu His Met Gly  
965 970 975

Met Ala Ala Arg Gln Leu Gly Glu His Ile Ala Thr Pro Val Phe Asp  
980 985 990

Gly Ala Asn Glu Glu Asp Val Trp Glu Thr Ile Lys Glu Ala Gly Met  
995 1000 1005

Asp Ala Asp Ala Lys Thr Val Leu Tyr Asp Gly Arg Thr Gly Glu Pro  
1010 1015 1020

Phe Asp Asn Lys Val Ser Val Gly Val Met Tyr Phe Ile Lys Leu Val  
1025 1030 1035 1040

His Met Val Asp Asp Lys Leu His Ala Arg Ser Thr Gly Pro Tyr Ser  
1045 1050 1055

Leu Val Thr Gln Gln Pro Leu Gly Gly Lys Ala Gln Phe Gly Gly Gln  
1060 1065 1070

Arg Phe Gly Glu Met Glu Val Trp Ala Leu Glu Ala Tyr Gly Ala Ser  
1075 1080 1085

Arg Thr Leu Gln Glu Ile Leu Thr Tyr Lys Ser Asp Asp Val Ile Gly  
 1090 1095 1100

Arg Val Asp Thr Tyr Glu Ala Ile Val Lys Gly Gln Arg Ile Pro Lys  
 1105 1110 1115 1120

Pro Gly Val Pro Glu Ser Phe Arg Val Leu Val Lys Glu Leu Gln Ser  
 1125 1130 1135

Leu Gly Leu Asp Leu Lys Val Leu Asp Lys Glu Gln Asn Glu Ile Asn  
 1140 1145 1150

Leu Lys Ala Glu Asp Asp Glu Ser Glu Asp Gln Val Val Asp Ser Leu  
 1155 1160 1165

Glu Glu Met Arg Lys Glu Gln Glu Glu Arg Arg Lys Glu Lys Glu  
 1170 1175 1180

Lys Glu Glu Pro Ser Thr Glu Ser  
 1185 1190

<210> 15  
 <211> 1407  
 <212> DNA  
 <213> *Alloioioccocus otitidis*

<220>  
 <221> CDS  
 <222> (25)..(1407)  
 <223>

<400> 15  
 aaagaccagg aaaggaagaa gacc ttg gca act aat att cat gaa gac cgc 51  
 Met Ala Thr Asn Ile His Glu Asp Arg  
 1 5

ctg cca cca caa aat att gaa gcg gag caa tcc gtc tta ggg tcc gtc 99  
 Leu Pro Pro Gln Asn Ile Glu Ala Glu Gln Ser Val Leu Gly Ser Val  
 10 15 20 25

ctc tta aat gca gaa gcc ttg gtg gcg gcc atg gaa tat gtg gat gaa 147  
 Leu Leu Asn Ala Glu Ala Leu Val Ala Ala Met Glu Tyr Val Asp Glu  
 30 35 40

gat gac ttt tac cgg cgg gcc cac cag ttg atc ttt aag gcc atg ata 195  
 Asp Asp Phe Tyr Arg Arg Ala His Gln Leu Ile Phe Lys Ala Met Ile  
 45 50 55

gac ctc tat gaa gac aac cag gcc att gat gtc att acc att aaa gac 243

Asp	Leu	Tyr	Glu	Asp	Asn	Gln	Ala	Ile	Asp	Val	Ile	Thr	Ile	Lys	Asp		
	60						65					70					
aag	ctg	gaa	gcc	aat	gac	cag	ttg	gag	gat	atc	ggg	ggt	gcc	tct	tac	291	
Lys	Leu	Glu	Ala	Asn	Asp	Gln	Leu	Glu	Asp	Ile	Gly	Gly	Ala	Ser	Tyr		
	75					80					85						
cta	gct	gag	att	gct	ggg	gtc	acc	cca	acc	gca	gct	aac	gtg	tcc	tat	339	
Leu	Ala	Glu	Ile	Ala	Gly	Val	Thr	Pro	Thr	Ala	Ala	Asn	Val	Ser	Tyr		
	90				95					100					105		
tac	gct	aag	att	gtg	gaa	gat	cgg	tct	ctt	ttg	cgc	aac	ttg	att	gcg	387	
Tyr	Ala	Lys	Ile	Val	Glu	Asp	Arg	Ser	Leu	Leu	Arg	Asn	Leu	Ile	Ala		
				110					115					120			
aca	gct	aat	gag	att	gcc	cag	tct	ggc	tac	gaa	gac	cat	gac	gat	gtg	435	
Thr	Ala	Asn	Glu	Ile	Ala	Gln	Ser	Gly	Tyr	Glu	Asp	His	Asp	Asp	Val		
			125					130					135				
cca	gaa	gtt	tta	aac	aat	gct	gag	cag	aag	atc	ttg	cag	gtt	tct	gaa	483	
Pro	Glu	Val	Leu	Asn	Asn	Ala	Glu	Gln	Lys	Ile	Leu	Gln	Val	Ser	Glu		
		140					145					150					
aaa	cga	aac	cgg	acc	ggc	ttt	gct	agt	att	tca	gaa	atc	ctc	cac	caa	531	
Lys	Arg	Asn	Arg	Thr	Gly	Phe	Ala	Ser	Ile	Ser	Glu	Ile	Leu	His	Gln		
	155					160					165						
acc	atc	gag	cat	att	gat	gaa	ctc	cac	caa	agg	gat	gaa	gag	atc	acc	579	
Thr	Ile	Glu	His	Ile	Asp	Glu	Leu	His	Gln	Arg	Asp	Glu	Glu	Ile	Thr		
	170				175					180					185		
ggg	att	tca	act	ggc	tac	ccc	tac	ctg	gac	agg	atg	act	tca	ggc	ctt	627	
Gly	Ile	Ser	Thr	Gly	Tyr	Pro	Tyr	Leu	Asp	Arg	Met	Thr	Ser	Gly	Leu		
			190						195					200			
cat	gaa	gat	gag	ttg	att	att	gtc	gca	gca	aga	ccg	ggt	gtg	ggg	aaa	675	
His	Glu	Asp	Glu	Leu	Ile	Ile	Val	Ala	Ala	Arg	Pro	Gly	Val	Gly	Lys		
			205					210					215				
acg	gct	ttt	gcc	ttg	aat	gtc	gcc	caa	aat	atc	ggg	aca	gcc	aca	gat	723	
Thr	Ala	Phe	Ala	Leu	Asn	Val	Ala	Gln	Asn	Ile	Gly	Thr	Ala	Thr	Asp		
		220					225					230					
gaa	act	att	gcg	att	ttt	tcc	ctt	gag	atg	ggg	gct	gaa	cag	ctg	gtc	771	
Glu	Thr	Ile	Ala	Ile	Phe	Ser	Leu	Glu	Met	Gly	Ala	Glu	Gln	Leu	Val		
	235					240					245						
aac	cgg	atg	tta	tgt	tca	gaa	ggc	agt	att	gat	gcc	act	aac	ctc	cga	819	
Asn	Arg	Met	Leu	Cys	Ser	Glu	Gly	Ser	Ile	Asp	Ala	Thr	Asn	Leu	Arg		
	250				255					260					265		
aat	ggc	aag	cta	acg	ccg	gaa	gaa	tat	gac	cgt	ttg	ttt	gtg	gcc	atg	867	
Asn	Gly	Lys	Leu	Thr	Pro	Glu	Glu	Tyr	Asp	Arg	Leu	Phe	Val	Ala	Met		
				270					275					280			
ggg	agc	ttg	tct	gaa	gct	gat	att	tat	att	gat	gac	act	ccc	ggc	atc	915	
Gly	Ser	Leu	Ser	Glu	Ala	Asp	Ile	Tyr	Ile	Asp	Asp	Thr	Pro	Gly	Ile		



285	290	295	
cgg aca gct gaa atc cgg gcc aag tgc cgc cgc ctg gtc caa gag aag			963
Arg Thr Ala Glu Ile Arg Ala Lys Cys Arg Arg Leu Val Gln Glu Lys			
300	305	310	
gga agt ctg ggc ttg att gtc att gac tac ctg caa ttg atc gaa gga			1011
Gly Ser Leu Gly Leu Ile Val Ile Asp Tyr Leu Gln Leu Ile Glu Gly			
315	320	325	
gct tca aac tat gaa tcc aga cag cag cag gtg tct gat ata tct cgg			1059
Ala Ser Asn Tyr Glu Ser Arg Gln Gln Val Ser Asp Ile Ser Arg			
330	335	340	345
cag ctg aag aag ctt tct aag gaa ctt tct gtc cca gtt att gcc ctg			1107
Gln Leu Lys Lys Leu Ser Lys Glu Leu Ser Val Pro Val Ile Ala Leu			
350	355	360	
tca caa ctg tcc cgg agt gtg gaa cag aga caa gac aag cgg ccc atc			1155
Ser Gln Leu Ser Arg Ser Val Glu Gln Arg Gln Asp Lys Arg Pro Ile			
365	370	375	
ctc agt gac ttg cgg gaa tca ggg tgc att gaa cag gat gcc gat att			1203
Leu Ser Asp Leu Arg Glu Ser Gly Ser Ile Glu Gln Asp Ala Asp Ile			
380	385	390	
gtg gcc ttc ctt tac cgg gag gac tac tac caa aat gaa gaa gat atc			1251
Val Ala Phe Leu Tyr Arg Glu Asp Tyr Tyr Gln Asn Glu Glu Asp Ile			
395	400	405	
gat gag gac ttt gtc gat aat agc gtg gaa gtc att atc gaa aaa aac			1299
Asp Glu Asp Phe Val Asp Asn Ser Val Glu Val Ile Ile Glu Lys Asn			
410	415	420	425
cgg tca gga gct cga gga aca gtc aag ttg aac ttt aag aaa gag ttc			1347
Arg Ser Gly Ala Arg Gly Thr Val Lys Leu Asn Phe Lys Lys Glu Phe			
430	435	440	
aac aaa ttt acc tcg att tct tac cgg tct gaa gat gaa gtc cca gcc			1395
Asn Lys Phe Thr Ser Ile Ser Tyr Arg Ser Glu Asp Glu Val Pro Ala			
445	450	455	
aac ttt ggc tag			1407
Asn Phe Gly			
460			

&lt;210&gt; 16

&lt;211&gt; 460

&lt;212&gt; PRT

&lt;213&gt; Alloiococcus otitidis

&lt;400&gt; 16

Met	Ala	Thr	Asn	Ile	His	Glu	Asp	Arg	Leu	Pro	Pro	Gln	Asn	Ile	Glu
1				5					10					15	

Ala Glu Gln Ser Val Leu Gly Ser Val Leu Leu Asn Ala Glu Ala Leu  
20 25 30

Val Ala Ala Met Glu Tyr Val Asp Glu Asp Asp Phe Tyr Arg Arg Ala  
35 40 45

His Gln Leu Ile Phe Lys Ala Met Ile Asp Leu Tyr Glu Asp Asn Gln  
50 55 60

Ala Ile Asp Val Ile Thr Ile Lys Asp Lys Leu Glu Ala Asn Asp Gln  
65 70 75 80

Leu Glu Asp Ile Gly Gly Ala Ser Tyr Leu Ala Glu Ile Ala Gly Val  
85 90 95

Thr Pro Thr Ala Ala Asn Val Ser Tyr Tyr Ala Lys Ile Val Glu Asp  
100 105 110

Arg Ser Leu Leu Arg Asn Leu Ile Ala Thr Ala Asn Glu Ile Ala Gln  
115 120 125

Ser Gly Tyr Glu Asp His Asp Asp Val Pro Glu Val Leu Asn Asn Ala  
130 135 140

Glu Gln Lys Ile Leu Gln Val Ser Glu Lys Arg Asn Arg Thr Gly Phe  
145 150 155 160

Ala Ser Ile Ser Glu Ile Leu His Gln Thr Ile Glu His Ile Asp Glu  
165 170 175

Leu His Gln Arg Asp Glu Glu Ile Thr Gly Ile Ser Thr Gly Tyr Pro  
180 185 190

Tyr Leu Asp Arg Met Thr Ser Gly Leu His Glu Asp Glu Leu Ile Ile  
195 200 205

Val Ala Ala Arg Pro Gly Val Gly Lys Thr Ala Phe Ala Leu Asn Val  
210 215 220

Ala Gln Asn Ile Gly Thr Ala Thr Asp Glu Thr Ile Ala Ile Phe Ser  
225 230 235 240

Leu Glu Met Gly Ala Glu Gln Leu Val Asn Arg Met Leu Cys Ser Glu

245 250 255

Gly Ser Ile Asp Ala Thr Asn Leu Arg Asn Gly Lys Leu Thr Pro Glu  
260 265 270

Glu Tyr Asp Arg Leu Phe Val Ala Met Gly Ser Leu Ser Glu Ala Asp  
275 280 285

Ile Tyr Ile Asp Asp Thr Pro Gly Ile Arg Thr Ala Glu Ile Arg Ala  
290 295 300

Lys Cys Arg Arg Leu Val Gln Glu Lys Gly Ser Leu Gly Leu Ile Val  
305 310 315 320

Ile Asp Tyr Leu Gln Leu Ile Glu Gly Ala Ser Asn Tyr Glu Ser Arg  
325 330 335

Gln Gln Gln Val Ser Asp Ile Ser Arg Gln Leu Lys Lys Leu Ser Lys  
340 345 350

Glu Leu Ser Val Pro Val Ile Ala Leu Ser Gln Leu Ser Arg Ser Val  
355 360 365

Glu Gln Arg Gln Asp Lys Arg Pro Ile Leu Ser Asp Leu Arg Glu Ser  
370 375 380

Gly Ser Ile Glu Gln Asp Ala Asp Ile Val Ala Phe Leu Tyr Arg Glu  
385 390 395 400

Asp Tyr Tyr Gln Asn Glu Glu Asp Ile Asp Glu Asp Phe Val Asp Asn  
405 410 415

Ser Val Glu Val Ile Ile Glu Lys Asn Arg Ser Gly Ala Arg Gly Thr  
420 425 430

Val Lys Leu Asn Phe Lys Lys Glu Phe Asn Lys Phe Thr Ser Ile Ser  
435 440 445

Tyr Arg Ser Glu Asp Glu Val Pro Ala Asn Phe Gly  
450 455 460

&lt;210&gt; 17

&lt;211&gt; 2484

&lt;212&gt; DNA

<213> *Alloioococcus otitidis*

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (10)..(2484)

&lt;223&gt;

&lt;400&gt; 17

```

aggaggaaa ttg ttt ttg gaa gag aga gat agc cgt tta gaa cag att aag      51
      Met Phe Leu Glu Glu Arg Asp Ser Arg Leu Glu Gln Ile Lys
        1              5              10

ctg tcc aag gag atg aaa aac tca ttc tta gac tat gcc atg agt gtc      99
Leu Ser Lys Glu Met Lys Asn Ser Phe Leu Asp Tyr Ala Met Ser Val
15              20              25              30

atc gtc tcc cgg gcc cta ccc gat gtc cgg gac ggc ttg aag ccg gtt      147
Ile Val Ser Arg Ala Leu Pro Asp Val Arg Asp Gly Leu Lys Pro Val
              35              40              45

cac cga aga atc ctg tac gga atg aat gaa ctg ggc tta acc ccg gac      195
His Arg Arg Ile Leu Tyr Gly Met Asn Glu Leu Gly Leu Thr Pro Asp
              50              55              60

aag tct tat aaa aag tct gcc cgt att gta ggg gat gtt atg ggg aaa      243
Lys Ser Tyr Lys Lys Ser Ala Arg Ile Val Gly Asp Val Met Gly Lys
              65              70              75

tac cac ccc cac ggt gac act gct att tat gac tcc atg gtc aga atg      291
Tyr His Pro His Gly Asp Thr Ala Ile Tyr Asp Ser Met Val Arg Met
              80              85              90

gcc cag gac ttt tct tac cga gtt ccc tta gtg gac ggc cat ggg aac      339
Ala Gln Asp Phe Ser Tyr Arg Val Pro Leu Val Asp Gly His Gly Asn
95              100              105              110

ttt ggg tcg gtt gac ggg gac gga gct gct gcc atg cgg tat acc gaa      387
Phe Gly Ser Val Asp Gly Asp Gly Ala Ala Ala Met Arg Tyr Thr Glu
              115              120              125

gcc cgg atg tcc aag atg gcc ttg gaa ctc ctg cga gac atc aac aag      435
Ala Arg Met Ser Lys Met Ala Leu Glu Leu Leu Arg Asp Ile Asn Lys
              130              135              140

gat acc att gac tac cac gat aac tat gat ggg act gag tcg gaa ccc      483
Asp Thr Ile Asp Tyr His Asp Asn Tyr Asp Gly Thr Glu Ser Glu Pro
              145              150              155

gat atc ctt cct gcc cgc ttc ccc aac ctc tta gtc aac ggg gct tcg      531
Asp Ile Leu Pro Ala Arg Phe Pro Asn Leu Leu Val Asn Gly Ala Ser
              160              165              170

ggg att gct gtt ggg atg gca acc aat atc cca ccc cac aat ctt aag      579
Gly Ile Ala Val Gly Met Ala Thr Asn Ile Pro Pro His Asn Leu Lys
              175              180              185              190

```

gaa gtc att gat gcc tgc gtc ctc ttg atg gaa aat gag gat gtg act	627
Glu Val Ile Asp Ala Cys Val Leu Leu Met Glu Asn Glu Asp Val Thr	
195 200 205	
gtg gct gac ctt atg gaa gtc tta cca gga cct gac ttt ccg act ggg	675
Val Ala Asp Leu Met Glu Val Leu Pro Gly Pro Asp Phe Pro Thr Gly	
210 215 220	
gct tcc ctt att ggt gtt tct ggc gtc cgc aag gct tat gag acc ggt	723
Ala Ser Leu Ile Gly Val Ser Gly Val Arg Lys Ala Tyr Glu Thr Gly	
225 230 235	
cgt ggg tcc att aaa tta cgg gcc aag tcc cgg atc gat gtc gac caa	771
Arg Gly Ser Ile Lys Leu Arg Ala Lys Ser Arg Ile Asp Val Asp Gln	
240 245 250	
aaa ggt aag gaa aga att att atc gac gaa att cct tac atg gtc aac	819
Lys Gly Lys Glu Arg Ile Ile Ile Asp Glu Ile Pro Tyr Met Val Asn	
255 260 265 270	
aag gcc aaa ttg gtc gaa aag att gcg gaa ctg gct cgg gac aag aaa	867
Lys Ala Lys Leu Val Glu Lys Ile Ala Glu Leu Ala Arg Asp Lys Lys	
275 280 285	
att gac ggc att acc gat tta aat gat gag tct gac cgg gaa ggc ttg	915
Ile Asp Gly Ile Thr Asp Leu Asn Asp Glu Ser Asp Arg Glu Gly Leu	
290 295 300	
cgg att gtg atc gat gta cgc cgg gat act tct gct ggt ata tta ctt	963
Arg Ile Val Ile Asp Val Arg Arg Asp Thr Ser Ala Gly Ile Leu Leu	
305 310 315	
aac aag ctt tac aaa atg acc caa ttg cag gtt tct ttt ggc ttt aac	1011
Asn Lys Leu Tyr Lys Met Thr Gln Leu Gln Val Ser Phe Gly Phe Asn	
320 325 330	
atg ctg gct atc gtc gat ggg gtg ccc aaa acc ttg ggc ctc aaa gac	1059
Met Leu Ala Ile Val Asp Gly Val Pro Lys Thr Leu Gly Leu Lys Asp	
335 340 345 350	
atc ctg acc cac tac tta gac cat caa aaa act gtt atc cgc agg cgg	1107
Ile Leu Thr His Tyr Leu Asp His Gln Lys Thr Val Ile Arg Arg Arg	
355 360 365	
aca gag ttt gac aag aac aag gct gaa tcg cgg gcc cac atc tta gaa	1155
Thr Glu Phe Asp Lys Asn Lys Ala Glu Ser Arg Ala His Ile Leu Glu	
370 375 380	
ggg ctt cgg act gcc tta gac cat atc gat gcc att att acc att atc	1203
Gly Leu Arg Thr Ala Leu Asp His Ile Asp Ala Ile Thr Thr Ile Ile	
385 390 395	
cgt cag tcc cag caa gct gaa gaa gcc aaa agt caa ttg atg gct tct	1251
Arg Gln Ser Gln Gln Ala Glu Glu Ala Lys Ser Gln Leu Met Ala Ser	
400 405 410	
tat gac ctc tct gac cgt caa gcc cag gcg att tta gac atg cgg atg	1299

Tyr Asp Leu Ser Asp Arg Gln Ala Gln Ala Ile Leu Asp Met Arg Met 415 420 425 430	
gtc cgg ttg act ggt ttg gaa aga gag aaa att gaa gat gaa tac gct Val Arg Leu Thr Gly Leu Glu Arg Glu Lys Ile Glu Asp Glu Tyr Ala 435 440 445	1347
gaa ctc tta gaa aaa atc gag gac ttg cgt gac atc ttg gcc cgg cca Glu Leu Leu Glu Lys Ile Glu Asp Leu Arg Asp Ile Leu Ala Arg Pro 450 455 460	1395
gaa cgg atc aag caa att atc aaa gaa gaa atg atc gaa att gct gaa Glu Arg Ile Lys Gln Ile Ile Lys Glu Glu Met Ile Glu Ile Ala Glu 465 470 475	1443
aaa cac ggc caa gac cgc cta act gac atc cgg gtt ggg gaa gag ttg Lys His Gly Gln Asp Arg Leu Thr Asp Ile Arg Val Gly Glu Glu Leu 480 485 490	1491
agc att gaa gac gaa gac ttg att gaa gaa gaa gat atc atc att acc Ser Ile Glu Asp Glu Asp Leu Ile Glu Glu Asp Ile Ile Ile Thr 495 500 505 510	1539
ctg tct cga aaa ggc tat atc aaa cgg atg ccg gct gga gaa ttc aag Leu Ser Arg Lys Gly Tyr Ile Lys Arg Met Pro Ala Gly Glu Phe Lys 515 520 525	1587
gcc caa aac cgc ggt ggc cgt ggg gtt aag ggg atg act acc aac gat Ala Gln Asn Arg Gly Gly Arg Gly Val Lys Gly Met Thr Thr Asn Asp 530 535 540	1635
ggg gac ttt gtt gaa cag ctg act ttc tgt tct agt cat gac caa atc Gly Asp Phe Val Glu Gln Leu Thr Phe Cys Ser Ser His Asp Gln Ile 545 550 555	1683
ctc ttc ttt acc aac caa ggc aag gtt tat aag atc aag gcc tac gaa Leu Phe Phe Thr Asn Gln Gly Lys Val Tyr Lys Ile Lys Ala Tyr Glu 560 565 570	1731
atc ccg gaa tat ggg cgt aat gcc aag gga att cct gcc atc aac ttt Ile Pro Glu Tyr Gly Arg Asn Ala Lys Gly Ile Pro Ala Ile Asn Phe 575 580 585 590	1779
tta aat ata gat aaa gat gaa tat att caa gcc atg gtc aac ttg act Leu Asn Ile Asp Lys Asp Glu Tyr Ile Gln Ala Met Val Asn Leu Thr 595 600 605	1827
gac cag gca gat gac cag gac caa ttc ttc ttt gcg aca aga ctt ggc Asp Gln Ala Asp Asp Gln Asp Gln Phe Phe Phe Ala Thr Arg Leu Gly 610 615 620	1875
cgg gtc aaa cgg acg gcc cag tct gaa ttt caa aat atc aga agt agc Arg Val Lys Arg Thr Ala Gln Ser Glu Phe Gln Asn Ile Arg Ser Ser 625 630 635	1923
ggg ttg aac gcg atc aat cta aat gaa ggc gat gaa ttg gtt aac gtg Gly Leu Asn Ala Ile Asn Leu Asn Glu Gly Asp Glu Leu Val Asn Val	1971

640	645	650	
gtc cct acc cac aat gac cag gcc att atc ctg gcc agc cag caa ggc			2019
Val Pro Thr His Asn Asp Gln Ala Ile Ile Leu Ala Ser Gln Gln Gly			
655	660	665	670
tat gcg gtc tac ttt gat gaa aaa gat atc cgt agc atg ggt cga ggg			2067
Tyr Ala Val Tyr Phe Asp Glu Lys Asp Ile Arg Ser Met Gly Arg Gly			
	675	680	685
gct gca ggt gtc cgt gga att cgc tta ggt gat ggc gac aca gtg gtt			2115
Ala Ala Gly Val Arg Gly Ile Arg Leu Gly Asp Gly Asp Thr Val Val			
	690	695	700
gcc atg gaa gtc tta gag ccg ggc caa gac gta tta gtc att act gaa			2163
Ala Met Glu Val Leu Glu Pro Gly Gln Asp Val Leu Val Ile Thr Glu			
	705	710	715
aaa ggg tac ggc aaa cga acc tcc caa gaa gag tac acc ctc cac aag			2211
Lys Gly Tyr Gly Lys Arg Thr Ser Gln Glu Glu Tyr Thr Leu His Lys			
	720	725	730
cga ggg ggc aag ggg gtt aaa acc ctt cat att acc gat aag aat ggt			2259
Arg Gly Gly Lys Gly Val Lys Thr Leu His Ile Thr Asp Lys Asn Gly			
	735	740	745
ccc cta att gga ctg aaa act gtc tct ggt ggt gag gac gtc atg att			2307
Pro Leu Ile Gly Leu Lys Thr Val Ser Gly Gly Glu Asp Val Met Ile			
	755	760	765
gtc acc gac caa ggt atc atg att cgt atc gaa gcc gac agc atc tct			2355
Val Thr Asp Gln Gly Ile Met Ile Arg Ile Glu Ala Asp Ser Ile Ser			
	770	775	780
cag acc tcc cgc cta acc caa ggt gtc cgt tta att cga ctt gaa gaa			2403
Gln Thr Ser Arg Leu Thr Gln Gly Val Arg Leu Ile Arg Leu Glu Glu			
	785	790	795
gat agc cgg gtg tca acg gta gcc ctc att gat att gac caa gag ctt			2451
Asp Ser Arg Val Ser Thr Val Ala Leu Ile Asp Ile Asp Gln Glu Leu			
	800	805	810
gac aat caa gtt aac caa acc gtt gag gaa taa			2484
Asp Asn Gln Val Asn Gln Thr Val Glu Glu			
	815	820	

&lt;210&gt; 18

&lt;211&gt; 824

&lt;212&gt; PRT

&lt;213&gt; Alloiococcus otitidis

&lt;400&gt; 18

Met	Phe	Leu	Glu	Glu	Arg	Asp	Ser	Arg	Leu	Glu	Gln	Ile	Lys	Leu	Ser
1				5					10				15		

Lys Glu Met Lys Asn Ser Phe Leu Asp Tyr Ala Met Ser Val Ile Val  
20 25 30

Ser Arg Ala Leu Pro Asp Val Arg Asp Gly Leu Lys Pro Val His Arg  
35 40 45

Arg Ile Leu Tyr Gly Met Asn Glu Leu Gly Leu Thr Pro Asp Lys Ser  
50 55 60

Tyr Lys Lys Ser Ala Arg Ile Val Gly Asp Val Met Gly Lys Tyr His  
65 70 75 80

Pro His Gly Asp Thr Ala Ile Tyr Asp Ser Met Val Arg Met Ala Gln  
85 90 95

Asp Phe Ser Tyr Arg Val Pro Leu Val Asp Gly His Gly Asn Phe Gly  
100 105 110

Ser Val Asp Gly Asp Gly Ala Ala Ala Met Arg Tyr Thr Glu Ala Arg  
115 120 125

Met Ser Lys Met Ala Leu Glu Leu Leu Arg Asp Ile Asn Lys Asp Thr  
130 135 140

Ile Asp Tyr His Asp Asn Tyr Asp Gly Thr Glu Ser Glu Pro Asp Ile  
145 150 155 160

Leu Pro Ala Arg Phe Pro Asn Leu Leu Val Asn Gly Ala Ser Gly Ile  
165 170 175

Ala Val Gly Met Ala Thr Asn Ile Pro Pro His Asn Leu Lys Glu Val  
180 185 190

Ile Asp Ala Cys Val Leu Leu Met Glu Asn Glu Asp Val Thr Val Ala  
195 200 205

Asp Leu Met Glu Val Leu Pro Gly Pro Asp Phe Pro Thr Gly Ala Ser  
210 215 220

Leu Ile Gly Val Ser Gly Val Arg Lys Ala Tyr Glu Thr Gly Arg Gly  
225 230 235 240

Ser Ile Lys Leu Arg Ala Lys Ser Arg Ile Asp Val Asp Gln Lys Gly



245 250 255

Lys Glu Arg Ile Ile Ile Asp Glu Ile Pro Tyr Met Val Asn Lys Ala  
260 265 270

Lys Leu Val Glu Lys Ile Ala Glu Leu Ala Arg Asp Lys Lys Ile Asp  
275 280 285

Gly Ile Thr Asp Leu Asn Asp Glu Ser Asp Arg Glu Gly Leu Arg Ile  
290 295 300

Val Ile Asp Val Arg Arg Asp Thr Ser Ala Gly Ile Leu Leu Asn Lys  
305 310 315 320

Leu Tyr Lys Met Thr Gln Leu Gln Val Ser Phe Gly Phe Asn Met Leu  
325 330 335

Ala Ile Val Asp Gly Val Pro Lys Thr Leu Gly Leu Lys Asp Ile Leu  
340 345 350

Thr His Tyr Leu Asp His Gln Lys Thr Val Ile Arg Arg Arg Thr Glu  
355 360 365

Phe Asp Lys Asn Lys Ala Glu Ser Arg Ala His Ile Leu Glu Gly Leu  
370 375 380

Arg Thr Ala Leu Asp His Ile Asp Ala Ile Ile Thr Ile Ile Arg Gln  
385 390 395 400

Ser Gln Gln Ala Glu Glu Ala Lys Ser Gln Leu Met Ala Ser Tyr Asp  
405 410 415

Leu Ser Asp Arg Gln Ala Gln Ala Ile Leu Asp Met Arg Met Val Arg  
420 425 430

Leu Thr Gly Leu Glu Arg Glu Lys Ile Glu Asp Glu Tyr Ala Glu Leu  
435 440 445

Leu Glu Lys Ile Glu Asp Leu Arg Asp Ile Leu Ala Arg Pro Glu Arg  
450 455 460

Ile Lys Gln Ile Ile Lys Glu Glu Met Ile Glu Ile Ala Glu Lys His  
465 470 475 480

Gly Gln Asp Arg Leu Thr Asp Ile Arg Val Gly Glu Glu Leu Ser Ile  
485 490 495

Glu Asp Glu Asp Leu Ile Glu Glu Glu Asp Ile Ile Ile Thr Leu Ser  
500 505 510

Arg Lys Gly Tyr Ile Lys Arg Met Pro Ala Gly Glu Phe Lys Ala Gln  
515 520 525

Asn Arg Gly Gly Arg Gly Val Lys Gly Met Thr Thr Asn Asp Gly Asp  
530 535 540

Phe Val Glu Gln Leu Thr Phe Cys Ser Ser His Asp Gln Ile Leu Phe  
545 550 555 560

Phe Thr Asn Gln Gly Lys Val Tyr Lys Ile Lys Ala Tyr Glu Ile Pro  
565 570 575

Glu Tyr Gly Arg Asn Ala Lys Gly Ile Pro Ala Ile Asn Phe Leu Asn  
580 585 590

Ile Asp Lys Asp Glu Tyr Ile Gln Ala Met Val Asn Leu Thr Asp Gln  
595 600 605

Ala Asp Asp Gln Asp Gln Phe Phe Phe Ala Thr Arg Leu Gly Arg Val  
610 615 620

Lys Arg Thr Ala Gln Ser Glu Phe Gln Asn Ile Arg Ser Ser Gly Leu  
625 630 635 640

Asn Ala Ile Asn Leu Asn Glu Gly Asp Glu Leu Val Asn Val Val Pro  
645 650 655

Thr His Asn Asp Gln Ala Ile Ile Leu Ala Ser Gln Gln Gly Tyr Ala  
660 665 670

Val Tyr Phe Asp Glu Lys Asp Ile Arg Ser Met Gly Arg Gly Ala Ala  
675 680 685

Gly Val Arg Gly Ile Arg Leu Gly Asp Gly Asp Thr Val Val Ala Met  
690 695 700

Glu Val Leu Glu Pro Gly Gln Asp Val Leu Val Ile Thr Glu Lys Gly  
705 710 715 720

Tyr Gly Lys Arg Thr Ser Gln Glu Glu Tyr Thr Leu His Lys Arg Gly  
725 730 735

Gly Lys Gly Val Lys Thr Leu His Ile Thr Asp Lys Asn Gly Pro Leu  
740 745 750

Ile Gly Leu Lys Thr Val Ser Gly Gly Glu Asp Val Met Ile Val Thr  
755 760 765

Asp Gln Gly Ile Met Ile Arg Ile Glu Ala Asp Ser Ile Ser Gln Thr  
770 775 780

Ser Arg Leu Thr Gln Gly Val Arg Leu Ile Arg Leu Glu Glu Asp Ser  
785 790 795 800

Arg Val Ser Thr Val Ala Leu Ile Asp Ile Asp Gln Glu Leu Asp Asn  
805 810 815

Gln Val Asn Gln Thr Val Glu Glu  
820

<210> 19  
<211> 1956  
<212> DNA  
<213> *Alloioococcus otitidis*

<220>  
<221> CDS  
<222> (7)..(1956)  
<223>

<400> 19  
cgtgta atg gct gaa gat aga cca tta aca cca aat gag tta gca gaa 48  
Met Ala Glu Asp Arg Pro Leu Thr Pro Asn Glu Leu Ala Glu  
1 5 10

ctg aaa aaa aca tat gat gct agt caa atc caa gtc tta gaa ggc cta 96  
Leu Lys Lys Thr Tyr Asp Ala Ser Gln Ile Gln Val Leu Glu Gly Leu  
15 20 25 30

gaa gca gtc aga gta cgg cgg ggt atg tac att ggg tcc acc agc aag 144  
Glu Ala Val Arg Val Arg Pro Gly Met Tyr Ile Gly Ser Thr Ser Lys  
35 40 45

gaa ggc ctc cac cac ttg gta tgg gag atc gtg gac aat gct att gac 192

Glu Gly Leu His His Leu Val Trp Glu Ile Val Asp Asn Ala Ile Asp	
50 55 60	
gaa gct atg gcc ggt tat gcc gac aag att tct gtt tcc att ttg gaa	240
Glu Ala Met Ala Gly Tyr Ala Asp Lys Ile Ser Val Ser Ile Leu Glu	
65 70 75	
ggc gac gtg atc caa gtg gct gat aac ggc cgg ggc atc ccg gtt gat	288
Gly Asp Val Ile Gln Val Ala Asp Asn Gly Arg Gly Ile Pro Val Asp	
80 85 90	
atc cag gaa aaa aca ggc cgg cca gct gtt gaa act gtc ttt aca gtc	336
Ile Gln Glu Lys Thr Gly Arg Pro Ala Val Glu Thr Val Phe Thr Val	
95 100 105 110	
ctc cac gct ggt ggg aaa ttt ggt ggc ggt ggt tac aag gtt tcc ggt	384
Leu His Ala Gly Lys Phe Gly Gly Gly Tyr Lys Val Ser Gly	
115 120 125	
ggc ctg cac ggt gta ggg tct tct gtg gtc aat gct ctc tcc gaa tac	432
Gly Leu His Gly Val Gly Ser Ser Val Val Asn Ala Leu Ser Glu Tyr	
130 135 140	
ctc caa gtc cag gtg cac cga gat ggt aaa atc tac caa caa gtt tac	480
Leu Gln Val Gln Val His Arg Asp Gly Lys Ile Tyr Gln Gln Val Tyr	
145 150 155	
aag cgg ggc ttg gtt gat tct gac ttg gaa gtg gtg ggt gag act gac	528
Lys Arg Gly Leu Val Asp Ser Asp Leu Glu Val Val Gly Glu Thr Asp	
160 165 170	
cac act gga act att gtt acc ttt aag gca gat agt ttg att ttt aaa	576
His Thr Gly Thr Ile Val Thr Phe Lys Ala Asp Ser Leu Ile Phe Lys	
175 180 185 190	
gac act act tct tat gac ttc aat acc tta gcc acc cgg atc cgg gag	624
Asp Thr Thr Ser Tyr Asp Phe Asn Thr Leu Ala Thr Arg Ile Arg Glu	
195 200 205	
ttg gcc ttc tta aac cga ggc ttg aat att tcc atc gaa gac aaa cgg	672
Leu Ala Phe Leu Asn Arg Gly Leu Asn Ile Ser Ile Glu Asp Lys Arg	
210 215 220	
caa gca ggc ggg cag tct ttg aac tac cac tat gaa ggt ggg ata tcg	720
Gln Ala Gly Gly Gln Ser Leu Asn Tyr His Tyr Glu Gly Gly Ile Ser	
225 230 235	
agt tat gtt gac cac ttg aat tcc agc cgt gaa gtt ctt tat gag acc	768
Ser Tyr Val Asp His Leu Asn Ser Ser Arg Glu Val Leu Tyr Glu Thr	
240 245 250	
cca att ttc ttg gaa ggg gaa gaa gaa ggg att tct gtg gaa att gcc	816
Pro Ile Phe Leu Glu Gly Glu Glu Glu Gly Ile Ser Val Glu Ile Ala	
255 260 265 270	
ctc cag cat acc gat agc ttc cat act aat tta atg agt ttt gcc aat	864
Leu Gln His Thr Asp Ser Phe His Thr Asn Leu Met Ser Phe Ala Asn	

275	280	285	
aac atc cac acc tat gag ggt ggc atg cat att tcc ggc ttc aag aca Asn Ile His Thr Tyr Glu Gly Gly Met His Ile Ser Gly Phe Lys Thr 290 295 300			912
gcc ctt acc cgg gcg gtc aac gac tat gcc cgg cag aat aac ttg ctc Ala Leu Thr Arg Ala Val Asn Asp Tyr Ala Arg Gln Asn Asn Leu Leu 305 310 315			960
cga gag tca gag gat aac ttt acc ggc gat gac gtt cgg gaa ggt ctg Arg Glu Ser Glu Asp Asn Phe Thr Gly Asp Asp Val Arg Glu Gly Leu 320 325 330			1008
acg gtg gtt ttg tca atc aag cac cca gac ccc caa ttt gaa gga caa Thr Val Val Leu Ser Ile Lys His Pro Asp Pro Gln Phe Glu Gly Gln 335 340 345 350			1056
acc aag act aag ctg ggg aac tct gaa gtc aga ggg ata att gac cgg Thr Lys Thr Lys Leu Gly Asn Ser Glu Val Arg Gly Ile Ile Asp Arg 355 360 365			1104
ctc ttt agc cag cac ttt gaa cgt tac ctc atg gaa aat cca aag gtt Leu Phe Ser Gln His Phe Glu Arg Tyr Leu Met Glu Asn Pro Lys Val 370 375 380			1152
ggc aag cgg att gtt gac aag gcg ctt ttg gct tcc aaa gcc cgc caa Gly Lys Arg Ile Val Asp Lys Ala Leu Leu Ala Ser Lys Ala Arg Gln 385 390 395			1200
gca gcc aag aga gcc cgg gaa gtc acc cgg aag aaa tca ggc tta gaa Ala Ala Lys Arg Ala Arg Glu Val Thr Arg Lys Lys Ser Gly Leu Glu 400 405 410			1248
att agc aac cta cca ggt aaa tta gct gac tgt tct tcc aaa gat cct Ile Ser Asn Leu Pro Gly Lys Leu Ala Asp Cys Ser Ser Lys Asp Pro 415 420 425 430			1296
gaa gaa tcc gaa ctc ttt att gta gaa ggg gat tca gct gga ggg tcg Glu Glu Ser Glu Leu Phe Ile Val Glu Gly Asp Ser Ala Gly Gly Ser 435 440 445			1344
gct aag caa ggt cgg tcc cgg gtt ttc cag gct att ttg ccg att cgt Ala Lys Gln Gly Arg Ser Arg Val Phe Gln Ala Ile Leu Pro Ile Arg 450 455 460			1392
ggc aag att ttg aat gtc gaa aaa gcc agc att gac cgt atc tta gcc Gly Lys Ile Leu Asn Val Glu Lys Ala Ser Ile Asp Arg Ile Leu Ala 465 470 475			1440
aat gaa gaa atc cgg tct ctc ttt aca gcc atg gga act ggc ttc ggg Asn Glu Glu Ile Arg Ser Leu Phe Thr Ala Met Gly Thr Gly Phe Gly 480 485 490			1488
gaa gaa ttt aat gtt gaa gaa gct cgc tac aat aag tta att atc atg Glu Glu Phe Asn Val Glu Glu Ala Arg Tyr Asn Lys Leu Ile Ile Met 495 500 505 510			1536

aca gat gct gat gtt gac gga gcc cac att cgg acc ttg ctc ttg acc 1584  
 Thr Asp Ala Asp Val Asp Gly Ala His Ile Arg Thr Leu Leu Leu Thr  
 515 520 525  
 ctt ctt tac cgg tat atg cgg ccc ttg att gaa gca ggt ttc gtc tac 1632  
 Leu Leu Tyr Arg Tyr Met Arg Pro Leu Ile Glu Ala Gly Phe Val Tyr  
 530 535 540  
 att gcc cag cca ccc ctc tac cag gtc aag caa ggc aag aag gtt aaa 1680  
 Ile Ala Gln Pro Pro Leu Tyr Gln Val Lys Gln Gly Lys Lys Val Lys  
 545 550 555  
 tac ttt gat agt gac cgg gaa ctg gac tcc tac ttg aaa gaa att cct 1728  
 Tyr Phe Asp Ser Asp Arg Glu Leu Asp Ser Tyr Leu Lys Glu Ile Pro  
 560 565 570  
 gac tca ccc aag cct tct gtc caa cgc tac aaa ggc tta gga gaa atg 1776  
 Asp Ser Pro Lys Pro Ser Val Gln Arg Tyr Lys Gly Leu Gly Glu Met  
 575 580 585 590  
 gat gct gag cag ttg tgg gaa acc acc atg aac cca gaa cac cgc cgc 1824  
 Asp Ala Glu Gln Leu Trp Glu Thr Thr Met Asn Pro Glu His Arg Arg  
 595 600 605  
 tta ctt cgg gta gac gta gac gac gcc att gag gct gac act att ttt 1872  
 Leu Leu Arg Val Asp Val Asp Ala Ile Glu Ala Asp Thr Ile Phe  
 610 615 620  
 gac atg ttg atg ggt gag gat gtc aaa ccc cgg cgc gac ttt atc aaa 1920  
 Asp Met Leu Met Gly Glu Asp Val Lys Pro Arg Arg Asp Phe Ile Lys  
 625 630 635  
 gaa aat gcc cgt tac gtg gaa aat atc gat atc tag 1956  
 Glu Asn Ala Arg Tyr Val Glu Asn Ile Asp Ile  
 640 645

&lt;210&gt; 20

&lt;211&gt; 649

&lt;212&gt; PRT

&lt;213&gt; Alloioococcus otitidis

&lt;400&gt; 20

Met Ala Glu Asp Arg Pro Leu Thr Pro Asn Glu Leu Ala Glu Leu Lys  
 1 5 10 15

Lys Thr Tyr Asp Ala Ser Gln Ile Gln Val Leu Glu Gly Leu Glu Ala  
 20 25 30

Val Arg Val Arg Pro Gly Met Tyr Ile Gly Ser Thr Ser Lys Glu Gly  
 35 40 45

Leu His His Leu Val Trp Glu Ile Val Asp Asn Ala Ile Asp Glu Ala

50

55

60

Met Ala Gly Tyr Ala Asp Lys Ile Ser Val Ser Ile Leu Glu Gly Asp  
65 70 75 80

Val Ile Gln Val Ala Asp Asn Gly Arg Gly Ile Pro Val Asp Ile Gln  
85 90 95

Glu Lys Thr Gly Arg Pro Ala Val Glu Thr Val Phe Thr Val Leu His  
100 105 110

Ala Gly Gly Lys Phe Gly Gly Gly Gly Tyr Lys Val Ser Gly Gly Leu  
115 120 125

His Gly Val Gly Ser Ser Val Val Asn Ala Leu Ser Glu Tyr Leu Gln  
130 135 140

Val Gln Val His Arg Asp Gly Lys Ile Tyr Gln Gln Val Tyr Lys Arg  
145 150 155 160

Gly Leu Val Asp Ser Asp Leu Glu Val Val Gly Glu Thr Asp His Thr  
165 170 175

Gly Thr Ile Val Thr Phe Lys Ala Asp Ser Leu Ile Phe Lys Asp Thr  
180 185 190

Thr Ser Tyr Asp Phe Asn Thr Leu Ala Thr Arg Ile Arg Glu Leu Ala  
195 200 205

Phe Leu Asn Arg Gly Leu Asn Ile Ser Ile Glu Asp Lys Arg Gln Ala  
210 215 220

Gly Gly Gln Ser Leu Asn Tyr His Tyr Glu Gly Gly Ile Ser Ser Tyr  
225 230 235 240

Val Asp His Leu Asn Ser Ser Arg Glu Val Leu Tyr Glu Thr Pro Ile  
245 250 255

Phe Leu Glu Gly Glu Glu Glu Gly Ile Ser Val Glu Ile Ala Leu Gln  
260 265 270

His Thr Asp Ser Phe His Thr Asn Leu Met Ser Phe Ala Asn Asn Ile  
275 280 285

His Thr Tyr Glu Gly Gly Met His Ile Ser Gly Phe Lys Thr Ala Leu  
290 295 300

Thr Arg Ala Val Asn Asp Tyr Ala Arg Gln Asn Asn Leu Leu Arg Glu  
305 310 315 320

Ser Glu Asp Asn Phe Thr Gly Asp Asp Val Arg Glu Gly Leu Thr Val  
325 330 335

Val Leu Ser Ile Lys His Pro Asp Pro Gln Phe Glu Gly Gln Thr Lys  
340 345 350

Thr Lys Leu Gly Asn Ser Glu Val Arg Gly Ile Ile Asp Arg Leu Phe  
355 360 365

Ser Gln His Phe Glu Arg Tyr Leu Met Glu Asn Pro Lys Val Gly Lys  
370 375 380

Arg Ile Val Asp Lys Ala Leu Leu Ala Ser Lys Ala Arg Gln Ala Ala  
385 390 395 400

Lys Arg Ala Arg Glu Val Thr Arg Lys Lys Ser Gly Leu Glu Ile Ser  
405 410 415

Asn Leu Pro Gly Lys Leu Ala Asp Cys Ser Ser Lys Asp Pro Glu Glu  
420 425 430

Ser Glu Leu Phe Ile Val Glu Gly Asp Ser Ala Gly Gly Ser Ala Lys  
435 440 445

Gln Gly Arg Ser Arg Val Phe Gln Ala Ile Leu Pro Ile Arg Gly Lys  
450 455 460

Ile Leu Asn Val Glu Lys Ala Ser Ile Asp Arg Ile Leu Ala Asn Glu  
465 470 475 480

Glu Ile Arg Ser Leu Phe Thr Ala Met Gly Thr Gly Phe Gly Glu Glu  
485 490 495

Phe Asn Val Glu Glu Ala Arg Tyr Asn Lys Leu Ile Ile Met Thr Asp  
500 505 510



Ala Asp Val Asp Gly Ala His Ile Arg Thr Leu Leu Leu Thr Leu Leu  
515 520 525

Tyr Arg Tyr Met Arg Pro Leu Ile Glu Ala Gly Phe Val Tyr Ile Ala  
530 535 540

Gln Pro Pro Leu Tyr Gln Val Lys Gln Gly Lys Lys Val Lys Tyr Phe  
545 550 555 560

Asp Ser Asp Arg Glu Leu Asp Ser Tyr Leu Lys Glu Ile Pro Asp Ser  
565 570 575

Pro Lys Pro Ser Val Gln Arg Tyr Lys Gly Leu Gly Glu Met Asp Ala  
580 585 590

Glu Gln Leu Trp Glu Thr Thr Met Asn Pro Glu His Arg Arg Leu Leu  
595 600 605

Arg Val Asp Val Asp Asp Ala Ile Glu Ala Asp Thr Ile Phe Asp Met  
610 615 620

Leu Met Gly Glu Asp Val Lys Pro Arg Arg Asp Phe Ile Lys Glu Asn  
625 630 635 640

Ala Arg Tyr Val Glu Asn Ile Asp Ile  
645

<210> 21

<211> 1218

<212> DNA

<213> Alloiococcus otitidis

<220>

<221> CDS

<222> (16)..(1218)

<223>

<400> 21

agacctaatac attttt ttg aaa tgg aga aag aca aaa acc atc tat ggt ata 51  
Met Lys Trp Arg Lys Thr Lys Thr Ile Tyr Gly Ile  
1 5 10

ctt aag aac aaa agg aag ttt gga ggg att ttt ttg aaa ttt tca gta 99  
Leu Lys Asn Lys Arg Lys Phe Gly Gly Ile Phe Leu Lys Phe Ser Val  
15 20 25

aaa cgg acg gaa ttt cta aaa gta tta aaa aaa gta cag att gca gtg 147

Lys Arg Thr Glu Phe Leu Lys Val Leu Lys Lys Val Gln Ile Ala Val	
30 35 40	
tct tct aaa agt acc atc gct atc ttg acc ggg att aaa tta gaa gcg	195
Ser Ser Lys Ser Thr Ile Ala Ile Leu Thr Gly Ile Lys Leu Glu Ala	
45 50 55 60	
gat aac cag ggt tta acc tta acc gga tct aac tcg gat atc tca gtt	243
Asp Asn Gln Gly Leu Thr Leu Thr Gly Ser Asn Ser Asp Ile Ser Val	
65 70 75	
gaa agt tac tta tct gtg acc gat gaa ggg gcg gat ttg gtt att gat	291
Glu Ser Tyr Leu Ser Val Thr Asp Glu Gly Ala Asp Leu Val Ile Asp	
80 85 90	
gag cgg ggg cag att gtc ttg caa cca gcc cgg tta ttt gcc aat atc	339
Glu Pro Gly Gln Ile Val Leu Gln Pro Ala Arg Leu Phe Ala Asn Ile	
95 100 105	
gtc caa aaa tta ccg gac acc cac ttt aag gta aac gtt agc caa ggc	387
Val Gln Lys Leu Pro Asp Thr His Phe Lys Val Asn Val Ser Gln Gly	
110 115 120	
cag caa acc caa atc acc tca gct tca gcc tcc ttt act atc aac ggc	435
Gln Gln Thr Gln Ile Thr Ser Ala Ser Ala Ser Phe Thr Ile Asn Gly	
125 130 135 140	
att gac gcc atg tcc tac ccc cac ttg cca gat atc gac ctg gag gaa	483
Ile Asp Ala Met Ser Tyr Pro His Leu Pro Asp Ile Asp Leu Glu Glu	
145 150 155	
tcc ttt acc ctg ccg gtt gac ctc ttt aaa aac atg atc aac cag act	531
Ser Phe Thr Leu Pro Val Asp Leu Phe Lys Asn Met Ile Asn Gln Thr	
160 165 170	
gtc atc gca gtc tcc aac cat gaa agt cgg ccc atc cta act ggg gtt	579
Val Ile Ala Val Ser Asn His Glu Ser Arg Pro Ile Leu Thr Gly Val	
175 180 185	
aac cta tct ctc aaa gag ggc cga ctc aag gca gtg gca acc gac agc	627
Asn Leu Ser Leu Lys Glu Gly Arg Leu Lys Ala Val Ala Thr Asp Ser	
190 195 200	
cac cgt ttg tcg caa cgg tcc atc caa tta gag tca gcg cct gat att	675
His Arg Leu Ser Gln Arg Ser Ile Gln Leu Glu Ser Ala Pro Asp Ile	
205 210 215 220	
tcc ttt gac att gtg ata cca ggt aag tct ttg act gaa ctg act aag	723
Ser Phe Asp Ile Val Ile Pro Gly Lys Ser Leu Thr Glu Leu Thr Lys	
225 230 235	
ttg atg gat gca gat gaa gaa gtc cgg gta gcc atc agc gac aac caa	771
Leu Met Asp Ala Asp Glu Glu Val Arg Val Ala Ile Ser Asp Asn Gln	
240 245 250	
atc cta ttt gcc ctc tcc agc agc cag ttt tac tct cgg ctc cta gaa	819
Ile Leu Phe Ala Leu Ser Ser Ser Gln Phe Tyr Ser Arg Leu Leu Glu	

255	260	265	
ggt aag tat cct gat acc gac cgc ttg atc cca ggc gac acc cca acg			867
Gly Lys Tyr Pro Asp Thr Asp Arg Leu Ile Pro Gly Asp Thr Pro Thr			
270	275	280	
gaa atc acc ttg gac acc aag gaa tta cag ggg gct gtt gac cgg gct			915
Glu Ile Thr Leu Asp Thr Lys Glu Leu Gln Gly Ala Val Asp Arg Ala			
285	290	295	300
tcc ctc ctc tcc cat gaa ggg aaa aac aat gtg gtc caa ctc aca gtg			963
Ser Leu Leu Ser His Glu Gly Lys Asn Asn Val Val Gln Leu Thr Val			
305	310		315
act gct gaa aag ttg gaa atc gaa ggc cag tca gct gaa gtg ggc cat			1011
Thr Ala Glu Lys Leu Glu Ile Glu Gly Gln Ser Ala Glu Val Gly His			
320	325		330
gtc caa gaa gaa att gac ttt ggc cac ttc caa ggc caa gac tta acc			1059
Val Gln Glu Glu Ile Asp Phe Gly His Phe Gln Gly Gln Asp Leu Thr			
335	340		345
att tct ttc aac ccc gac tac tta aaa gag gcc ttg gct acc ttt ggt			1107
Ile Ser Phe Asn Pro Asp Tyr Leu Lys Glu Ala Leu Ala Thr Phe Gly			
350	355		360
caa gga gaa att aag ttg aaa tta gtt tcg acc ttg cga ccc ttt gtc			1155
Gln Gly Glu Ile Lys Leu Lys Leu Val Ser Thr Leu Arg Pro Phe Val			
365	370		380
atc gtc cca agt gag gac caa gga gac ttt atc caa ctt att act cca			1203
Ile Val Pro Ser Glu Asp Gln Gly Asp Phe Ile Gln Leu Ile Thr Pro			
385	390		395
atc cga aca gcc taa			1218
Ile Arg Thr Ala			
400			

&lt;210&gt; 22

&lt;211&gt; 400

&lt;212&gt; PRT

&lt;213&gt; Alloiococcus otitidis

&lt;400&gt; 22

Met	Lys	Trp	Arg	Lys	Thr	Lys	Thr	Ile	Tyr	Gly	Ile	Leu	Lys	Asn	Lys
1				5					10					15	

Arg	Lys	Phe	Gly	Gly	Ile	Phe	Leu	Lys	Phe	Ser	Val	Lys	Arg	Thr	Glu
			20					25						30	

Phe	Leu	Lys	Val	Leu	Lys	Lys	Val	Gln	Ile	Ala	Val	Ser	Ser	Lys	Ser
			35				40							45	

Thr Ile Ala Ile Leu Thr Gly Ile Lys Leu Glu Ala Asp Asn Gln Gly  
50 55 60

Leu Thr Leu Thr Gly Ser Asn Ser Asp Ile Ser Val Glu Ser Tyr Leu  
65 70 75 80

Ser Val Thr Asp Glu Gly Ala Asp Leu Val Ile Asp Glu Pro Gly Gln  
85 90 95

Ile Val Leu Gln Pro Ala Arg Leu Phe Ala Asn Ile Val Gln Lys Leu  
100 105 110

Pro Asp Thr His Phe Lys Val Asn Val Ser Gln Gly Gln Gln Thr Gln  
115 120 125

Ile Thr Ser Ala Ser Ala Ser Phe Thr Ile Asn Gly Ile Asp Ala Met  
130 135 140

Ser Tyr Pro His Leu Pro Asp Ile Asp Leu Glu Glu Ser Phe Thr Leu  
145 150 155 160

Pro Val Asp Leu Phe Lys Asn Met Ile Asn Gln Thr Val Ile Ala Val  
165 170 175

Ser Asn His Glu Ser Arg Pro Ile Leu Thr Gly Val Asn Leu Ser Leu  
180 185 190

Lys Glu Gly Arg Leu Lys Ala Val Ala Thr Asp Ser His Arg Leu Ser  
195 200 205

Gln Arg Ser Ile Gln Leu Glu Ser Ala Pro Asp Ile Ser Phe Asp Ile  
210 215 220

Val Ile Pro Gly Lys Ser Leu Thr Glu Leu Thr Lys Leu Met Asp Ala  
225 230 235 240

Asp Glu Glu Val Arg Val Ala Ile Ser Asp Asn Gln Ile Leu Phe Ala  
245 250 255

Leu Ser Ser Ser Gln Phe Tyr Ser Arg Leu Leu Glu Gly Lys Tyr Pro  
260 265 270

Asp Thr Asp Arg Leu Ile Pro Gly Asp Thr Pro Thr Glu Ile Thr Leu

275

280

285

Asp Thr Lys Glu Leu Gln Gly Ala Val Asp Arg Ala Ser Leu Leu Ser  
 290 295 300

His Glu Gly Lys Asn Asn Val Val Gln Leu Thr Val Thr Ala Glu Lys  
 305 310 315 320

Leu Glu Ile Glu Gly Gln Ser Ala Glu Val Gly His Val Gln Glu Glu  
 325 330 335

Ile Asp Phe Gly His Phe Gln Gly Gln Asp Leu Thr Ile Ser Phe Asn  
 340 345 350

Pro Asp Tyr Leu Lys Glu Ala Leu Ala Thr Phe Gly Gln Gly Glu Ile  
 355 360 365

Lys Leu Lys Leu Val Ser Thr Leu Arg Pro Phe Val Ile Val Pro Ser  
 370 375 380

Glu Asp Gln Gly Asp Phe Ile Gln Leu Ile Thr Pro Ile Arg Thr Ala  
 385 390 395 400

&lt;210&gt; 23

&lt;211&gt; 1317

&lt;212&gt; DNA

<213> *Alloiococcus otitidis*

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (25)..(1317)

&lt;223&gt;

&lt;400&gt; 23

tcaataactg ctttttttagg agct ttg cag atg aat tgg aaa gaa acc atc 51  
 Met Gln Met Asn Trp Lys Glu Thr Ile  
 1 5

agt ctc atc aac acc acc cgg ggg acc gga gac aag aaa aat ttg aac 99  
 Ser Leu Ile Asn Thr Thr Arg Gly Thr Gly Asp Lys Lys Asn Leu Asn  
 10 15 20 25

cgg atg cga ctt tta ctc aaa gag cta ggt aat cct gaa aca gac ttg 147  
 Arg Met Arg Leu Leu Leu Lys Glu Leu Gly Asn Pro Glu Thr Asp Leu  
 30 35 40

ccg gtc atc cac gtt gct ggc acc aat ggc aaa ggg acg acc tgt gct Pro Val Ile His Val Ala Gly Thr Asn Gly Lys Gly Thr Thr Cys Ala 45 50 55	195
tat att gcc cac agc ttg gcc cgt gct ggt tat aaa aca gga ctt tac Tyr Ile Ala His Ser Leu Ala Arg Ala Gly Tyr Lys Thr Gly Leu Tyr 60 65 70	243
acc agc ccc cac ctg gag cgg gtc aat gaa cgg atc cgg att aat gac Thr Ser Pro His Leu Glu Arg Val Asn Glu Arg Ile Arg Ile Asn Asp 75 80 85	291
cgc tac ata tcc gac caa gac tta atg gct ttg acc ggt caa att gcc Arg Tyr Ile Ser Asp Gln Asp Leu Met Ala Leu Thr Gly Gln Ile Ala 90 95 100 105	339
ccc atc att gac cat cta gaa gac tgc ttg ggt gag aaa tac tat tct Pro Ile Ile Asp His Leu Glu Asp Cys Leu Gly Glu Lys Tyr Tyr Ser 110 115 120	387
ttt gaa att tta act gcc ctt gcc ttc ttg tac ttc cag caa gca ggg Phe Glu Ile Leu Thr Ala Leu Ala Phe Leu Tyr Phe Gln Gln Ala Gly 125 130 135	435
gtg gac ttt tta gtt tta gaa act ggg gta ggg gga aaa att gat gcg Val Asp Phe Leu Val Leu Glu Thr Gly Val Gly Gly Lys Ile Asp Ala 140 145 150	483
acc aat gtg gtg ccc gct cca ctg gtc tca gtc att atc tct att ggc Thr Asn Val Val Pro Ala Pro Leu Val Ser Val Ile Ile Ser Ile Gly 155 160 165	531
tat gac cac acc cat gtc ttg ggt aat acc ctg gaa gac att acc cgg Tyr Asp His Thr His Val Leu Gly Asn Thr Leu Glu Asp Ile Thr Arg 170 175 180 185	579
cac aag gca ggg att att aag aaa ggc tgt ccg gtg gtg gtg ggc cct His Lys Ala Gly Ile Ile Lys Lys Gly Cys Pro Val Val Val Gly Pro 190 195 200	627
ctt gcc gac cat tta ttg gct att gtt aaa gag gtg tcc aaa gaa atg Leu Ala Asp His Leu Leu Ala Ile Val Lys Glu Val Ser Lys Glu Met 205 210 215	675
gac agt aat tta acc att gtc cat ccc gac aag ttt gac att gtt cat Asp Ser Asn Leu Thr Ile Val His Pro Asp Lys Phe Asp Ile Val His 220 225 230	723
caa acc ctt gac tac cag tcc ttt aaa tac ggt ggg gac ttg gtt tta Gln Thr Leu Asp Tyr Gln Ser Phe Lys Tyr Gly Gly Asp Leu Val Leu 235 240 245	771
gag act caa atg att ggt aac cac cag ctg gta aac act gcc cta gct Glu Thr Gln Met Ile Gly Asn His Gln Leu Val Asn Thr Ala Leu Ala 250 255 260 265	819
tat gaa gcc ttg aag att gtc caa caa tct tac ccc gat ttg aca gat	867

Tyr Glu Ala Leu Lys Ile Val Gln Gln Ser Tyr Pro Asp Leu Thr Asp  
 270 275 280  
 tta gat ata tta gaa ggc ttg aag acg acc cac tgg cca ggc cgg atg 915  
 Leu Asp Ile Leu Glu Gly Leu Lys Thr Thr His Trp Pro Gly Arg Met  
 285 290 295  
 caa aag cta tct gac cag cca gtg gtt gtt ctt gat ggg gcc cac aac 963  
 Gln Lys Leu Ser Asp Gln Pro Val Val Val Leu Asp Gly Ala His Asn  
 300 305 310  
 gaa atc ggg gtc aag gct ctt aga cag tca att gac cac ttt ttc ccc 1011  
 Glu Ile Gly Val Lys Ala Leu Arg Gln Ser Ile Asp His Phe Phe Pro  
 315 320 325  
 ggc aaa aaa atc acc tat ttt gcc gga atg atg gtc gaa aaa gac ttc 1059  
 Gly Lys Lys Ile Thr Tyr Phe Ala Gly Met Met Val Glu Lys Asp Phe  
 330 335 340 345  
 gcc aaa atg ttt gac ctc ctg ggg gaa aca gct gat aaa ttt tac ttg 1107  
 Ala Lys Met Phe Asp Leu Leu Gly Glu Thr Ala Asp Lys Phe Tyr Leu  
 350 355 360  
 att tca ccc gat ttg act cgc ggt ttt gat gtc gac caa gcc gtt caa 1155  
 Ile Ser Pro Asp Leu Thr Arg Gly Phe Asp Val Asp Gln Ala Val Gln  
 365 370 375  
 tct ttg act gac aag ggc tac cag gct tcc agt gtg gct agc ctc caa 1203  
 Ser Leu Thr Asp Lys Gly Tyr Gln Ala Ser Ser Val Ala Ser Leu Gln  
 380 385 390  
 gcc atc tta gac tac ata aac cag caa gca aaa gca gat gaa att atc 1251  
 Ala Ile Leu Asp Tyr Ile Asn Gln Gln Ala Lys Ala Asp Glu Ile Ile  
 395 400 405  
 att atc ttt ggc tcc ctc tac ttg gtt ggc gac ttc cta aaa ctt tac 1299  
 Ile Ile Phe Gly Ser Leu Tyr Leu Val Gly Asp Phe Leu Lys Leu Tyr  
 410 415 420 425  
 cat gaa gca tcc ggt taa 1317  
 His Glu Ala Ser Gly  
 430

&lt;210&gt; 24

&lt;211&gt; 430

&lt;212&gt; PRT

&lt;213&gt; Alloiococcus otitidis

&lt;400&gt; 24

Met Gln Met Asn Trp Lys Glu Thr Ile Ser Leu Ile Asn Thr Thr Arg  
 1 5 10 15

Gly Thr Gly Asp Lys Lys Asn Leu Asn Arg Met Arg Leu Leu Leu Lys  
 20 25 30

Glu Leu Gly Asn Pro Glu Thr Asp Leu Pro Val Ile His Val Ala Gly  
35 40 45

Thr Asn Gly Lys Gly Thr Thr Cys Ala Tyr Ile Ala His Ser Leu Ala  
50 55 60

Arg Ala Gly Tyr Lys Thr Gly Leu Tyr Thr Ser Pro His Leu Glu Arg  
65 70 75 80

Val Asn Glu Arg Ile Arg Ile Asn Asp Arg Tyr Ile Ser Asp Gln Asp  
85 90 95

Leu Met Ala Leu Thr Gly Gln Ile Ala Pro Ile Ile Asp His Leu Glu  
100 105 110

Asp Cys Leu Gly Glu Lys Tyr Tyr Ser Phe Glu Ile Leu Thr Ala Leu  
115 120 125

Ala Phe Leu Tyr Phe Gln Gln Ala Gly Val Asp Phe Leu Val Leu Glu  
130 135 140

Thr Gly Val Gly Gly Lys Ile Asp Ala Thr Asn Val Val Pro Ala Pro  
145 150 155 160

Leu Val Ser Val Ile Ile Ser Ile Gly Tyr Asp His Thr His Val Leu  
165 170 175

Gly Asn Thr Leu Glu Asp Ile Thr Arg His Lys Ala Gly Ile Ile Lys  
180 185 190

Lys Gly Cys Pro Val Val Val Gly Pro Leu Ala Asp His Leu Leu Ala  
195 200 205

Ile Val Lys Glu Val Ser Lys Glu Met Asp Ser Asn Leu Thr Ile Val  
210 215 220

His Pro Asp Lys Phe Asp Ile Val His Gln Thr Leu Asp Tyr Gln Ser  
225 230 235 240

Phe Lys Tyr Gly Gly Asp Leu Val Leu Glu Thr Gln Met Ile Gly Asn  
245 250 255



His Gln Leu Val Asn Thr Ala Leu Ala Tyr Glu Ala Leu Lys Ile Val  
                   260                  265                  270

Gln Gln Ser Tyr Pro Asp Leu Thr Asp Leu Asp Ile Leu Glu Gly Leu  
                   275                  280                  285

Lys Thr Thr His Trp Pro Gly Arg Met Gln Lys Leu Ser Asp Gln Pro  
                   290                  295                  300

Val Val Val Leu Asp Gly Ala His Asn Glu Ile Gly Val Lys Ala Leu  
                   305                  310                  315                  320

Arg Gln Ser Ile Asp His Phe Phe Pro Gly Lys Lys Ile Thr Tyr Phe  
                   325                  330                  335

Ala Gly Met Met Val Glu Lys Asp Phe Ala Lys Met Phe Asp Leu Leu  
                   340                  345                  350

Gly Glu Thr Ala Asp Lys Phe Tyr Leu Ile Ser Pro Asp Leu Thr Arg  
                   355                  360                  365

Gly Phe Asp Val Asp Gln Ala Val Gln Ser Leu Thr Asp Lys Gly Tyr  
                   370                  375                  380

Gln Ala Ser Ser Val Ala Ser Leu Gln Ala Ile Leu Asp Tyr Ile Asn  
                   385                  390                  395                  400

Gln Gln Ala Lys Ala Asp Glu Ile Ile Ile Ile Phe Gly Ser Leu Tyr  
                   405                  410                  415

Leu Val Gly Asp Phe Leu Lys Leu Tyr His Glu Ala Ser Gly  
                   420                  425                  430

<210> 25

<211> 1653

<212> DNA

<213> *Alloiococcus otitidis*

<220>

<221> CDS

<222> (91)..(1653)

<223>

<400> 25

cttctgtgtt catttttaaat cttatctgaa acaaatgatt tttcaactct tttttatctt

60

acttttaaag ttttagttag gagccttagc ttg tac cgt atc tct atg aaa gac	114
Met Tyr Arg Ile Ser Met Lys Asp	
1 5	
ttg cat gcc cta tta gct agc aag cag cag ttg aaa gaa gtg gtc ggt	162
Leu His Ala Leu Leu Ala Ser Lys Gln Gln Leu Lys Glu Val Val Gly	
10 15 20	
ccc gac caa gtt tgg cat tac aat ttg cct caa ggg gaa ttg gcc gac	210
Pro Asp Gln Val Trp His Tyr Asn Leu Pro Gln Gly Glu Leu Ala Asp	
25 30 35 40	
caa gtt ttt gac aaa ctt tcc tac aat tcc caa act gcc tcc tca gac	258
Gln Val Phe Asp Lys Leu Ser Tyr Asn Ser Gln Thr Ala Ser Ser Asp	
45 50 55	
acc ctt ttc ttt tgc aag ggt gct tcc ttt aaa aga gac tac cta gcc	306
Thr Leu Phe Phe Cys Lys Gly Ala Ser Phe Lys Arg Asp Tyr Leu Ala	
60 65 70	
cag gcg gtt gac cag ggt gtc caa gtc tat att tcc gaa aaa ttg tat	354
Gln Ala Val Asp Gln Gly Val Gln Val Tyr Ile Ser Glu Lys Leu Tyr	
75 80 85	
caa ggc ctg gat gct tat gcc atc att gtc cgt gac atc cgc cag acc	402
Gln Gly Leu Asp Ala Tyr Ala Ile Ile Val Arg Asp Ile Arg Gln Thr	
90 95 100	
atg gcc cta gtc gct aag gct ttt tac cag gct cca gat gaa aaa ttg	450
Met Ala Leu Val Ala Lys Ala Phe Tyr Gln Ala Pro Asp Glu Lys Leu	
105 110 115 120	
acc ctg att ggc att acc ggg acc aag ggc aag aca acc aca agt tac	498
Thr Leu Ile Gly Ile Thr Gly Thr Lys Gly Lys Thr Thr Thr Ser Tyr	
125 130 135	
ctc ctc aaa tcc atc ctg gac cag gac caa gcc ggt aag aca gct att	546
Leu Leu Lys Ser Ile Leu Asp Gln Asp Gln Ala Gly Lys Thr Ala Ile	
140 145 150	
att tca acc ttg ggg att tcc tta gac ggc cag acc caa gaa gaa gcc	594
Ile Ser Thr Leu Gly Ile Ser Leu Asp Gly Gln Thr Gln Glu Glu Ala	
155 160 165	
tcc ctg acc act cct gaa gcc ttg gac ctc tac cag atg att gcc cgg	642
Ser Leu Thr Thr Pro Glu Ala Leu Asp Leu Tyr Gln Met Ile Ala Arg	
170 175 180	
gcc caa gac cag ggg atg gac caa ttg att atg gaa gta tct agc caa	690
Ala Gln Asp Gln Gly Met Asp Gln Leu Ile Met Glu Val Ser Ser Gln	
185 190 195 200	
gcc tac aag atg gac cgg gtc tat gga ctg act ttc gac ttt gga gcc	738
Ala Tyr Lys Met Asp Arg Val Tyr Gly Leu Thr Phe Asp Phe Gly Ala	
205 210 215	

ttc tta aat att tcg cct gac cat atc ggc cct aat gag cac cca gat Phe Leu Asn Ile Ser Pro Asp His Ile Gly Pro Asn Glu His Pro Asp 220 225 230	786
atg gaa gat tac ttc tat tgt aaa agt cgt ttg gtt aaa cat tcc aag Met Glu Asp Tyr Phe Tyr Cys Lys Ser Arg Leu Val Lys His Ser Lys 235 240 245	834
ttg gcc ttg ctc aac gct gga ctt gac cag cta gac tac tta aaa gac Leu Ala Leu Leu Asn Ala Gly Leu Asp Gln Leu Asp Tyr Leu Lys Asp 250 255 260	882
ctt agc caa aaa aat ggc ggt cag gtc caa gtt tac ggc caa gat ccc Leu Ser Gln Lys Asn Gly Gly Gln Val Gln Val Tyr Gly Gln Asp Pro 265 270 275 280	930
aag act tgt gac tac tat ttt gag gtt aac aac cag gac agc cgc cgc Lys Thr Cys Asp Tyr Tyr Phe Glu Val Asn Asn Gln Asp Ser Arg Arg 285 290 295	978
ttt gcc att aaa agc caa agc cct gat gac ttg gcc att gat ggg gat Phe Ala Ile Lys Ser Gln Ser Pro Asp Asp Leu Ala Ile Asp Gly Asp 300 305 310	1026
tac caa ttt gaa atg ttg ggt gat ttt aac aag gag aat gcc ctt tgt Tyr Gln Phe Glu Met Leu Gly Asp Phe Asn Lys Glu Asn Ala Leu Cys 315 320 325	1074
gcc gct ctt ata gcg ggg cat tta gaa gtt ggg caa gag gcc att tac Ala Ala Leu Ile Ala Gly His Leu Glu Val Gly Gln Glu Ala Ile Tyr 330 335 340	1122
caa gga ata gcc cag gcc caa gtg cca gga cgg atg cag cat tat act Gln Gly Ile Ala Gln Ala Gln Val Pro Gly Arg Met Gln His Tyr Thr 345 350 355 360	1170
tat ggc aac aat cac atc tat gta gac ttt gcc cac aat tac atc agc Tyr Gly Asn Asn His Ile Tyr Val Asp Phe Ala His Asn Tyr Ile Ser 365 370 375	1218
ttg aaa aat ctt ttt gat ttt gcc caa gac caa cac ccg gac cac acc Leu Lys Asn Leu Phe Asp Phe Ala Gln Asp Gln His Pro Asp His Thr 380 385 390	1266
atg gtg gtt gtc ttg ggg gcc cct ggc aac aag ggg gtg tct cgc cgc Met Val Val Val Leu Gly Ala Pro Gly Asn Lys Gly Val Ser Arg Arg 395 400 405	1314
aag gat atg gga tac ttg ctg tcc caa tac caa ggg gaa gtt atc ttg Lys Asp Met Gly Tyr Leu Leu Ser Gln Tyr Gln Gly Glu Val Ile Leu 410 415 420	1362
acc gaa gat gac ccc aat ttt gaa gac gtt caa gct atc tgc caa gaa Thr Glu Asp Asp Pro Asn Phe Glu Asp Val Gln Ala Ile Cys Gln Glu 425 430 435 440	1410

att gcc caa tac att gat ggc ccc atc cag gtg acc ttt aat gat aac 1458  
 Ile Ala Gln Tyr Ile Asp Gly Pro Ile Gln Val Thr Phe Asn Asp Asn  
                   445                                  450                                  455  
  
 cgg ata aat gcc atc caa gac ctc cta gag tcc tta acc cca gaa agt 1506  
 Arg Ile Asn Ala Ile Gln Asp Leu Leu Glu Ser Leu Thr Pro Glu Ser  
                   460                                  465                                  470  
  
 caa aaa gtc atc ctg ctt gca ggc aag ggg tcc gac cag tac atg ctg 1554  
 Gln Lys Val Ile Leu Leu Ala Gly Lys Gly Ser Asp Gln Tyr Met Leu  
                   475                                  480                                  485  
  
 cgg cgg ggt gtg aag gaa gat tat gcg gga gac cac aaa ttg gtt gaa 1602  
 Arg Arg Gly Val Lys Glu Asp Tyr Ala Gly Asp His Lys Leu Val Glu  
                   490                                  495                                  500  
  
 gca ttt tta aac cag caa aag act tct tct cat gag aag ctt gag ggt 1650  
 Ala Phe Leu Asn Gln Gln Lys Thr Ser Ser His Glu Lys Leu Glu Gly  
 505                                  510                                  515                                  520  
  
 tag. 1653

&lt;210&gt; 26

&lt;211&gt; 520

&lt;212&gt; PRT

&lt;213&gt; Alloiococcus otitidis

&lt;400&gt; 26

Met Tyr Arg Ile Ser Met Lys Asp Leu His Ala Leu Leu Ala Ser Lys  
 1                  5                                  10                                  15

Gln Gln Leu Lys Glu Val Val Gly Pro Asp Gln Val Trp His Tyr Asn  
                   20                                  25                                  30

Leu Pro Gln Gly Glu Leu Ala Asp Gln Val Phe Asp Lys Leu Ser Tyr  
                   35                                  40                                  45

Asn Ser Gln Thr Ala Ser Ser Asp Thr Leu Phe Phe Cys Lys Gly Ala  
 50                                  55                                  60

Ser Phe Lys Arg Asp Tyr Leu Ala Gln Ala Val Asp Gln Gly Val Gln  
 65                                  70                                  75                                  80

Val Tyr Ile Ser Glu Lys Leu Tyr Gln Gly Leu Asp Ala Tyr Ala Ile  
                   85                                  90                                  95

Ile Val Arg Asp Ile Arg Gln Thr Met Ala Leu Val Ala Lys Ala Phe  
                   100                                  105                                  110

Tyr Gln Ala Pro Asp Glu Lys Leu Thr Leu Ile Gly Ile Thr Gly Thr  
115 120 125

Lys Gly Lys Thr Thr Thr Ser Tyr Leu Leu Lys Ser Ile Leu Asp Gln  
130 135 140

Asp Gln Ala Gly Lys Thr Ala Ile Ile Ser Thr Leu Gly Ile Ser Leu  
145 150 155 160

Asp Gly Gln Thr Gln Glu Glu Ala Ser Leu Thr Thr Pro Glu Ala Leu  
165 170 175

Asp Leu Tyr Gln Met Ile Ala Arg Ala Gln Asp Gln Gly Met Asp Gln  
180 185 190

Leu Ile Met Glu Val Ser Ser Gln Ala Tyr Lys Met Asp Arg Val Tyr  
195 200 205

Gly Leu Thr Phe Asp Phe Gly Ala Phe Leu Asn Ile Ser Pro Asp His  
210 215 220

Ile Gly Pro Asn Glu His Pro Asp Met Glu Asp Tyr Phe Tyr Cys Lys  
225 230 235 240

Ser Arg Leu Val Lys His Ser Lys Leu Ala Leu Leu Asn Ala Gly Leu  
245 250 255

Asp Gln Leu Asp Tyr Leu Lys Asp Leu Ser Gln Lys Asn Gly Gly Gln  
260 265 270

Val Gln Val Tyr Gly Gln Asp Pro Lys Thr Cys Asp Tyr Tyr Phe Glu  
275 280 285

Val Asn Asn Gln Asp Ser Arg Arg Phe Ala Ile Lys Ser Gln Ser Pro  
290 295 300

Asp Asp Leu Ala Ile Asp Gly Asp Tyr Gln Phe Glu Met Leu Gly Asp  
305 310 315 320

Phe Asn Lys Glu Asn Ala Leu Cys Ala Ala Leu Ile Ala Gly His Leu  
325 330 335

Glu Val Gly Gln Glu Ala Ile Tyr Gln Gly Ile Ala Gln Ala Gln Val

340

345

350

Pro Gly Arg Met Gln His Tyr Thr Tyr Gly Asn Asn His Ile Tyr Val  
355 360 365

Asp Phe Ala His Asn Tyr Ile Ser Leu Lys Asn Leu Phe Asp Phe Ala  
370 375 380

Gln Asp Gln His Pro Asp His Thr Met Val Val Val Leu Gly Ala Pro  
385 390 395 400

Gly Asn Lys Gly Val Ser Arg Arg Lys Asp Met Gly Tyr Leu Leu Ser  
405 410 415

Gln Tyr Gln Gly Glu Val Ile Leu Thr Glu Asp Asp Pro Asn Phe Glu  
420 425 430

Asp Val Gln Ala Ile Cys Gln Glu Ile Ala Gln Tyr Ile Asp Gly Pro  
435 440 445

Ile Gln Val Thr Phe Asn Asp Asn Arg Ile Asn Ala Ile Gln Asp Leu  
450 455 460

Leu Glu Ser Leu Thr Pro Glu Ser Gln Lys Val Ile Leu Leu Ala Gly  
465 470 475 480

Lys Gly Ser Asp Gln Tyr Met Leu Arg Arg Gly Val Lys Glu Asp Tyr  
485 490 495

Ala Gly Asp His Lys Leu Val Glu Ala Phe Leu Asn Gln Gln Lys Thr  
500 505 510

Ser Ser His Glu Lys Leu Glu Gly  
515 520

&lt;210&gt; 27

&lt;211&gt; 636

&lt;212&gt; DNA

<213> *Alloiococcus otitidis*

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (25)..(636)

&lt;223&gt;

<400> 27  
aggactgaaa ggaaggttgg agtc atg aag tgg ttg agt cgg atc ttg att 51  
Met Lys Trp Leu Ser Arg Ile Leu Ile  
1 5

ggt ggt ggg ata ggc ttt ctg att gcc ttt ggc tac acg att tat gac 99  
Val Val Gly Ile Gly Phe Leu Ile Ala Phe Gly Tyr Thr Ile Tyr Asp  
10 15 20 25

cat gct aac tcg aca tcg gtt acc cta gaa gaa gcc cag gtg gcc ctg 147  
His Ala Asn Ser Thr Ser Val Thr Leu Glu Ala Gln Val Ala Leu  
30 35 40

gaa gaa agc cgg gcc cag gct gct gaa gct ggg gac ggg gac cag gat 195  
Glu Glu Ser Arg Ala Gln Ala Ala Glu Ala Gly Asp Gly Asp Gln Asp  
45 50 55

ggc caa gat ggg gcg agt gac atc gat atc caa aac tac cag cct gaa 243  
Gly Gln Asp Gly Ala Ser Asp Ile Asp Ile Gln Asn Tyr Gln Pro Glu  
60 65 70

gct ggg gag gct ttt ggg gtc tta gat att ccc aaa ctc gac cgg agc 291  
Ala Gly Glu Ala Phe Gly Val Leu Asp Ile Pro Lys Leu Asp Arg Ser  
75 80 85

att ggc att gta gcc gga acg gat gca gac tct ctt aag aag ggg gta 339  
Ile Gly Ile Val Ala Gly Thr Asp Ala Asp Ser Leu Lys Lys Gly Val  
90 95 100 105

ggt cac gtt gag aat aca gtc ttc cct ggc caa ggc gaa caa att gtc 387  
Gly His Val Glu Asn Thr Val Phe Pro Gly Gln Gly Glu Gln Ile Val  
110 115 120

ctc tct ggc cac cgg gat acc gtc ttc cgg gac ttt ggc gaa tta gaa 435  
Leu Ser Gly His Arg Asp Thr Val Phe Arg Asp Phe Gly Glu Leu Glu  
125 130 135

att ggc gac aat ttt atc gtt caa atg cct tac ggg gac tat gaa tat 483  
Ile Gly Asp Asn Phe Ile Val Gln Met Pro Tyr Gly Asp Tyr Glu Tyr  
140 145 150

gag att cag gac tat gaa att gtc gac cgg gat gat acc tcc gtc atc 531  
Glu Ile Gln Asp Tyr Glu Ile Val Asp Arg Asp Asp Thr Ser Val Ile  
155 160 165

cgg cct atg ggg gaa gaa gtc tta gtg gtt tca acc tgc tac ccc ttt 579  
Arg Pro Met Gly Glu Glu Val Leu Val Val Ser Thr Cys Tyr Pro Phe  
170 175 180 185

gaa ttt tac ggt ttt gcc cct gac cgc ttt gtt ttc tat tgt tac ccc 627  
Glu Phe Tyr Gly Phe Ala Pro Asp Arg Phe Val Phe Tyr Cys Tyr Pro  
190 195 200

ggt gaa taa 636  
Val Glu

<210> 28  
 <211> 203  
 <212> PRT  
 <213> *Alloioicoccus otitidis*

<400> 28  
 Met Lys Trp Leu Ser Arg Ile Leu Ile Val Val Gly Ile Gly Phe Leu  
 1 5 10 15

Ile Ala Phe Gly Tyr Thr Ile Tyr Asp His Ala Asn Ser Thr Ser Val  
 20 25 30

Thr Leu Glu Glu Ala Gln Val Ala Leu Glu Glu Ser Arg Ala Gln Ala  
 35 40 45

Ala Glu Ala Gly Asp Gly Asp Gln Asp Gly Gln Asp Gly Ala Ser Asp  
 50 55 60

Ile Asp Ile Gln Asn Tyr Gln Pro Glu Ala Gly Glu Ala Phe Gly Val  
 65 70 75 80

Leu Asp Ile Pro Lys Leu Asp Arg Ser Ile Gly Ile Val Ala Gly Thr  
 85 90 95

Asp Ala Asp Ser Leu Lys Lys Gly Val Gly His Val Glu Asn Thr Val  
 100 105 110

Phe Pro Gly Gln Gly Glu Gln Ile Val Leu Ser Gly His Arg Asp Thr  
 115 120 125

Val Phe Arg Asp Phe Gly Glu Leu Glu Ile Gly Asp Asn Phe Ile Val  
 130 135 140

Gln Met Pro Tyr Gly Asp Tyr Glu Tyr Glu Ile Gln Asp Tyr Glu Ile  
 145 150 155 160

Val Asp Arg Asp Asp Thr Ser Val Ile Arg Pro Met Gly Glu Glu Val  
 165 170 175

Leu Val Val Ser Thr Cys Tyr Pro Phe Glu Phe Tyr Gly Phe Ala Pro  
 180 185 190

Asp Arg Phe Val Phe Tyr Cys Tyr Pro Val Glu  
 195 200



<210> 29  
 <211> 1290  
 <212> DNA  
 <213> *Alloioiococcus otitidis*

<220>  
 <221> CDS  
 <222> (1)..(1290)  
 <223>

<400> 29  
 atg cag tat gca gaa ctt ctt gac ctc ctg ccc cta caa gaa caa ggg 48  
 Met Gln Tyr Ala Glu Leu Leu Asp Leu Leu Pro Leu Gln Glu Gln Gly  
 1 5 10 15  
 aag atg gat ttg ggg cta gca acc atg acc cag gtg atg gac tta ttg 96  
 Lys Met Asp Leu Gly Leu Ala Thr Met Thr Gln Val Met Asp Leu Leu  
 20 25 30  
 ggc aag ccc caa gac cag gtc ccc atg gtt cat atc gct ggc acc aat 144  
 Gly Lys Pro Gln Asp Gln Val Pro Met Val His Ile Ala Gly Thr Asn  
 35 40 45  
 ggc aag ggg tcg gcc gca gcc ttt aca gag cga ata ctc agg gag gct 192  
 Gly Lys Gly Ser Ala Ala Ala Phe Thr Glu Arg Ile Leu Arg Glu Ala  
 50 55 60  
 ggc tac aag gtc ggc ttg tat att tcc cct tcc cta gtg gaa ttt aat 240  
 Gly Tyr Lys Val Gly Leu Tyr Ile Ser Pro Ser Leu Val Glu Phe Asn  
 65 70 75 80  
 gaa cgg atc caa atc aat ggc caa gcc aca agt gat gat cag ttg ctc 288  
 Glu Arg Ile Gln Ile Asn Gly Gln Ala Thr Ser Asp Asp Gln Leu Leu  
 85 90 95  
 aag gca gtc aag acc cta agc cag gcc tta gaa ggc aca tcc ctt tgc 336  
 Lys Ala Val Lys Thr Leu Ser Gln Ala Leu Glu Gly Thr Ser Leu Cys  
 100 105 110  
 ctg act gaa ttt gaa ctt ttt act gcc ctg gcc ttt ttg acc ttc cag 384  
 Leu Thr Glu Phe Glu Leu Phe Thr Ala Leu Ala Phe Leu Thr Phe Gln  
 115 120 125  
 gac cag gct tgt gat ata gcc gtt gta gag gtc gga tta gga gga cgg 432  
 Asp Gln Ala Cys Asp Ile Ala Val Val Glu Val Gly Leu Gly Gly Arg  
 130 135 140  
 tta gat gct acc aat gtg ata agc cgt cct gcc gtc acc gcc att acc 480  
 Leu Asp Ala Thr Asn Val Ile Ser Arg Pro Ala Val Thr Ala Ile Thr  
 145 150 155 160  
 aag att ggc atg gac cat acc gct ttt tta ggg gat agc ctg cca gaa 528  
 Lys Ile Gly Met Asp His Thr Ala Phe Leu Gly Asp Ser Leu Pro Glu  
 165 170 175  
 ata gcc ggt gag aag gca gcc atc gcc aaa gcc ggc tcg cct atg gtg 576

Ile Ala Gly	Glu Lys Ala Ala	Ile Ala Lys Ala Gly	Ser Pro Met Val	
180		185	190	
gtc tat ccc	cag ggg cca gaa	gtg act cgg gtg atc	caa aat cag gcg	624
Val Tyr Pro	Gln Gly Pro Glu	Val Thr Arg Val Ile	Gln Asn Gln Ala	
195		200	205	
gac cgg gta	gga gcc tct ctg	acc cta att tct	caa tcc gac ctg act	672
Asp Arg Val	Gly Ala Ser Leu	Thr Leu Ile Ser	Gln Ser Asp Leu Thr	
210		215	220	
tat aac ctg	act tcg gac ctc	ttg caa gac ttt	gaa tac aag cag gtt	720
Tyr Asn Leu	Thr Ser Asp Leu	Leu Gln Asp Phe	Glu Tyr Lys Gln Val	
225		230	235	240
ccc tac cgc	att tca ctt tta	gaa gat tat caa	att tac aac gcc ctg	768
Pro Tyr Arg	Ile Ser Leu Leu	Glu Asp Tyr Gln	Ile Tyr Asn Ala Leu	
	245	250	255	
gta gca ctc	gaa atc tct ttt	gcc tta cag gat	gct ggc tgg cag att	816
Val Ala Leu	Glu Ile Ser Phe	Ala Leu Gln Asp	Ala Gly Trp Gln Ile	
	260	265	270	
agc cct aaa	gcc att aaa caa	ggg ttg gtt gag	acc cgc tgg ccc ggc	864
Ser Pro Lys	Ala Ile Lys Gln	Gly Leu Val Glu	Thr Arg Trp Pro Gly	
	275	280	285	
cgt ttt gaa	ctt atc gcc tct	cat ccg acc gtg	atc gtt gat ggg tct	912
Arg Phe Glu	Leu Ile Ala Ser	His Pro Thr Val	Ile Val Asp Gly Ser	
	290	295	300	
cat aat gaa	gac ggc ctg cag	gct ctc ttg gct	aac cta gac cgc tac	960
His Asn Glu	Asp Gly Leu Gln	Ala Leu Leu Ala	Asn Leu Asp Arg Tyr	
305		310	315	320
ttt cca gaa	caa aaa agg att	ggg atc gta ggc	atg ttg gcc gac aag	1008
Phe Pro Glu	Gln Lys Arg Ile	Gly Ile Val Gly	Met Leu Ala Asp Lys	
	325	330	335	
gat gtt gat	gcc gcc cta gct	cct tta acc aaa	agc ttt gac cgg ctt	1056
Asp Val Asp	Ala Ala Leu Ala	Pro Leu Thr Lys	Ser Phe Asp Arg Leu	
	340	345	350	
tat acg gtg	aca ccc gat tcg	ccg cgg ggg atg	gca gcc cct caa atg	1104
Tyr Thr Val	Thr Pro Asp Ser	Pro Arg Gly Met	Ala Ala Pro Gln Met	
	355	360	365	
aaa gaa aaa	ctg acc gaa atg	gtg tcg ccg tct	act cgg gtc ata gct	1152
Lys Glu Lys	Leu Thr Glu Met	Val Ser Pro Ser	Thr Arg Val Ile Ala	
	370	375	380	
tgt gaa agt	tat aac cag gcc	tta gac ctg gca	ggg caa gta gcc ggc	1200
Cys Glu Ser	Tyr Asn Gln Ala	Leu Asp Leu Ala	Gly Gln Val Ala Gly	
385		390	395	400
gga gat gac	cta att gtc gtt	ttt gga agt ttt	tat att gtt ggt aag	1248
Gly Asp Asp	Leu Ile Val Val	Phe Gly Ser Phe	Tyr Ile Val Gly Lys	

Lys Ile Gly Met Asp His Thr Ala Phe Leu Gly Asp Ser Leu Pro Glu  
165 170 175

Ile Ala Gly Glu Lys Ala Ala Ile Ala Lys Ala Gly Ser Pro Met Val  
 180 185 190

Val Tyr Pro Gln Gly Pro Glu Val Thr Arg Val Ile Gln Asn Gln Ala  
 195 200 205

Asp Arg Val Gly Ala Ser Leu Thr Leu Ile Ser Gln Ser Asp Leu Thr  
 210 215 220

Tyr Asn Leu Thr Ser Asp Leu Leu Gln Asp Phe Glu Tyr Lys Gln Val  
 225 230 235 240

Pro Tyr Arg Ile Ser Leu Leu Glu Asp Tyr Gln Ile Tyr Asn Ala Leu  
 245 250 255

Val Ala Leu Glu Ile Ser Phe Ala Leu Gln Asp Ala Gly Trp Gln Ile  
 260 265 270

Ser Pro Lys Ala Ile Lys Gln Gly Leu Val Glu Thr Arg Trp Pro Gly  
 275 280 285

Arg Phe Glu Leu Ile Ala Ser His Pro Thr Val Ile Val Asp Gly Ser  
 290 295 300

His Asn Glu Asp Gly Leu Gln Ala Leu Leu Ala Asn Leu Asp Arg Tyr  
 305 310 315 320

Phe Pro Glu Gln Lys Arg Ile Gly Ile Val Gly Met Leu Ala Asp Lys  
 325 330 335

Asp Val Asp Ala Ala Leu Ala Pro Leu Thr Lys Ser Phe Asp Arg Leu  
 340 345 350

Tyr Thr Val Thr Pro Asp Ser Pro Arg Gly Met Ala Ala Pro Gln Met  
 355 360 365

Lys Glu Lys Leu Thr Glu Met Val Ser Pro Ser Thr Arg Val Ile Ala  
 370 375 380

Cys Glu Ser Tyr Asn Gln Ala Leu Asp Leu Ala Gly Gln Val Ala Gly  
 385 390 395 400

Gly Asp Asp Leu Ile Val Val Phe Gly Ser Phe Tyr Ile Val Gly Lys

405

410

415

Phe Arg Gln Leu Ile Leu Ala Arg Arg Asn Gly Glu Val  
 420 425

&lt;210&gt; 31

&lt;211&gt; 387

&lt;212&gt; DNA

<213> *Alloioococcus otitidis*

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (16)..(387)

&lt;223&gt;

&lt;400&gt; 31

agaaagagga ataag atg gat aaa aga gat aag ata cgc ttg caa ggg atg 51  
 Met Asp Lys Arg Asp Lys Ile Arg Leu Gln Gly Met  
 1 5 10

act ttt cac ggc cac cac ggt ttg atg gag gcc gaa acc aag ttg ggt 99  
 Thr Phe His Gly His His Gly Leu Met Glu Ala Glu Thr Lys Leu Gly  
 15 20 25

cag att ttt aaa gtc gac ctt gtc tta gta act gac ctc aag tta gcg 147  
 Gln Ile Phe Lys Val Asp Leu Val Leu Val Thr Asp Leu Lys Leu Ala  
 30 35 40

ggg caa aca gac aag atg ggg cac agt atc cac tac ggg gaa gtt tat 195  
 Gly Gln Thr Asp Lys Met Gly His Ser Ile His Tyr Gly Glu Val Tyr  
 45 50 55 60

gac ctg gtc aag tcc att gtg gaa ggt acc ccc ttt aag ctt ttg gag 243  
 Asp Leu Val Lys Ser Ile Val Glu Gly Thr Pro Phe Lys Leu Leu Glu  
 65 70 75

tcc ttg gcg gaa acc cta gcc caa gaa gtt ctc aag act ttt gac cag 291  
 Ser Leu Ala Glu Thr Leu Ala Gln Glu Val Leu Lys Thr Phe Asp Gln  
 80 85 90

gtt gag gag gtc ttg gtc cgg gtc aac aaa ccc cag gcc ccg att cct 339  
 Val Glu Glu Val Leu Val Arg Val Asn Lys Pro Gln Ala Pro Ile Pro  
 95 100 105

ggg gtc ttt gac aat gta gcg gtg gaa atc acc cgg gcc cgt cac tag 387  
 Gly Val Phe Asp Asn Val Ala Val Glu Ile Thr Arg Ala Arg His  
 110 115 120

&lt;210&gt; 32

&lt;211&gt; 123

&lt;212&gt; PRT

<213> *Alloioococcus otitidis*

&lt;400&gt; 32

Met Asp Lys Arg Asp Lys Ile Arg Leu Gln Gly Met Thr Phe His Gly  
1 5 10 15

His His Gly Leu Met Glu Ala Glu Thr Lys Leu Gly Gln Ile Phe Lys  
20 25 30

Val Asp Leu Val Leu Val Thr Asp Leu Lys Leu Ala Gly Gln Thr Asp  
35 40 45

Lys Met Gly His Ser Ile His Tyr Gly Glu Val Tyr Asp Leu Val Lys  
50 55 60

Ser Ile Val Glu Gly Thr Pro Phe Lys Leu Leu Glu Ser Leu Ala Glu  
65 70 75 80

Thr Leu Ala Gln Glu Val Leu Lys Thr Phe Asp Gln Val Glu Glu Val  
85 90 95

Leu Val Arg Val Asn Lys Pro Gln Ala Pro Ile Pro Gly Val Phe Asp  
100 105 110

Asn Val Ala Val Glu Ile Thr Arg Ala Arg His  
115 120

<210> 33

<211> 552

<212> DNA

<213> *Alloioicoccus otitidis*

<220>

<221> CDS

<222> (22) .. (552)

<223>

<400> 33

ataggttaagg aggaatatag a gtg aag ggt gtt atg ata gga ctc ggt tct 51  
Met Lys Gly Val Met Ile Gly Leu Gly Ser  
1 5 10

aat atg ggg act aag ttg gct tac tta aac cgg gct ttg gcc aaa ata 99  
Asn Met Gly Thr Lys Leu Ala Tyr Leu Asn Arg Ala Leu Ala Lys Ile  
15 20 25

aat agc cta gac cag gta gca gtc aag caa gtt tca aag gtt tac cag 147  
Asn Ser Leu Asp Gln Val Ala Val Lys Gln Val Ser Lys Val Tyr Gln  
30 35 40

act gaa ccg gtg ggc tac aag gac cag gac gat ttt tac aat atg gtt 195  
Thr Glu Pro Val Gly Tyr Lys Asp Gln Asp Asp Phe Tyr Asn Met Val

Lys Asp Gln Asp Asp Phe Tyr Asn Met Val Ala Gly Leu Glu Ile Glu  
50 55 60

Pro Gly Lys Thr Pro Leu Asp Leu Leu Glu Asp Leu Leu Ala Ile Glu  
65 70 75 80

Ala Asp Leu Asp Arg Lys Arg Thr Ile Lys Asn Gly Pro Arg Thr Ile  
85 90 95

Asp Leu Asp Val Leu Leu Val Glu Gly Gln Glu Ile Asp His Pro Lys  
100 105 110

Leu Gln Val Pro His Pro Arg Leu Gln Asp Arg Ala Phe Val Leu Val  
115 120 125

Pro Leu Ala Glu Leu Asp Pro Asn Tyr Leu Val Pro Gly Ile Asp Lys  
130 135 140

Thr Val Ala Asp Leu Leu Ala Ser Leu Asn Gln Thr Asp Leu Ala Gly  
145 150 155 160

Val Glu Ala Leu Gly Gln Leu Thr Asn Leu Leu Glu Asp Arg Glu Ala  
165 170 175

<210> 35

<211> 1242

<212> DNA

<213> *Alloioococcus otitidis*

<220>

<221> CDS

<222> (40)..(1242)

<223>

<400> 35

aatcttcttta atatcgcttg gcccaagacc gctataata gtg gta agt gat tat 54  
Met Val Ser Asp Tyr  
1 5

ttt agg agg ttc aat atg caa ata gga att gac aag ctg gct ttt gcg 102  
Phe Arg Arg Phe Asn Met Gln Ile Gly Ile Asp Lys Leu Ala Phe Ala  
10 15 20

act cca acc agg tac ttg gaa atg gcg agt ctg gcc caa gcc cgg tcc 150  
Thr Pro Thr Arg Tyr Leu Glu Met Ala Ser Leu Ala Gln Ala Arg Ser  
25 30 35

caa gac cct aat aaa tat atc aag ggg cta ggc caa gaa gcc atg gct 198  
Gln Asp Pro Asn Lys Tyr Ile Lys Gly Leu Gly Gln Glu Ala Met Ala  
40 45 50



gtc cct gaa gaa agt gat gat gcc gtc agc ttg gcg gct aat gcc ggt	246
Val Pro Glu Glu Ser Asp Asp Ala Val Ser Leu Ala Ala Asn Ala Gly	
55 60 65	
aat tta atc tta agt gaa gaa gac aag gct gct att gac atg gtg ata	294
Asn Leu Ile Leu Ser Glu Glu Asp Lys Ala Ala Ile Asp Met Val Ile	
70 75 80 85	
gtc ggt acc gaa tct ggg gtc gac cag tcc aag tcg gca gcc agc tgg	342
Val Gly Thr Glu Ser Gly Val Asp Gln Ser Lys Ser Ala Ala Ser Trp	
90 95 100	
gtt cat gac ctg ttg ggg atc aac ccc cat gct aga agc ctg gag atc	390
Val His Asp Leu Leu Gly Ile Asn Pro His Ala Arg Ser Leu Glu Ile	
105 110 115	
aag caa gcc tgc tac ggg gct acg gct gga ctc aaa cta gct gtg gcc	438
Lys Gln Ala Cys Tyr Gly Ala Thr Ala Gly Leu Lys Leu Ala Val Ala	
120 125 130	
cac cta gcc tta aac cct gac tcc aag gtt tta gtc atc ggt tca gac	486
His Leu Ala Leu Asn Pro Asp Ser Lys Val Leu Val Ile Gly Ser Asp	
135 140 145	
ata gcc aag tat ggt ttg gaa aca ggg ggc gag ccc act caa gga gct	534
Ile Ala Lys Tyr Gly Leu Glu Thr Gly Gly Glu Pro Thr Gln Gly Ala	
150 155 160 165	
ggg gcg gtc gcc atc tta gtc agc cgt gac cct gca att gct gtg gtc	582
Gly Ala Val Ala Ile Leu Val Ser Arg Asp Pro Ala Ile Ala Val Val	
170 175 180	
aac aat gac agt gcc atg ctg acc aaa aat att gca gac ttt tgg cga	630
Asn Asn Asp Ser Ala Met Leu Thr Lys Asn Ile Ala Asp Phe Trp Arg	
185 190 195	
ccc aac tac agc gat tat gcc cat gta gat ggc aag ttc tcc aac cag	678
Pro Asn Tyr Ser Asp Tyr Ala His Val Asp Gly Lys Phe Ser Asn Gln	
200 205 210	
gca tac ttg tcc aac cta gca gaa gtc tgg cgc cag tat aag atc aaa	726
Ala Tyr Leu Ser Asn Leu Ala Glu Val Trp Arg Gln Tyr Lys Ile Lys	
215 220 225	
aac cag ctg tct gct aag gat ttc aag gcc atg gtc ttc cac agc ccc	774
Asn Gln Leu Ser Ala Lys Asp Phe Lys Ala Met Val Phe His Ser Pro	
230 235 240 245	
tat acc aag atg ggg aaa aag gcc tta ctc aaa cta gga gat tat gaa	822
Tyr Thr Lys Met Gly Lys Lys Ala Leu Leu Lys Leu Gly Asp Tyr Glu	
250 255 260	
gac cag aaa gag att gac cgc ttg ctg gcc tat tac gag cct ggt cgc	870
Asp Gln Lys Glu Ile Asp Arg Leu Leu Ala Tyr Tyr Glu Pro Gly Arg	
265 270 275	

tac tac aat aag cgg gtc ggt aat atc tat act ggg tct ctt tac ttg 918  
 Tyr Tyr Asn Lys Arg Val Gly Asn Ile Tyr Thr Gly Ser Leu Tyr Leu  
 280 285 290  
 agt ttg att tcc ctc tta gac cag gta agt gac ctg gag gct ggc gac 966  
 Ser Leu Ile Ser Leu Leu Asp Gln Val Ser Asp Leu Glu Ala Gly Asp  
 295 300 305  
 cgg att ggc ctc tat tct tat ggg tct ggt gcc gtt gga gag ttc ttt 1014  
 Arg Ile Gly Leu Tyr Ser Tyr Gly Ser Gly Ala Val Gly Glu Phe Phe  
 310 315 320 325  
 agc att cgg ctc cag cca ggt tac aag gaa agc tta cag caa gtt gac 1062  
 Ser Ile Arg Leu Gln Pro Gly Tyr Lys Glu Ser Leu Gln Gln Val Asp  
 330 335 340  
 ttc gac cag gtt gtc aac cag cgt tca gca tta gag atg tac agc tat 1110  
 Phe Asp Gln Val Val Asn Gln Arg Ser Ala Leu Glu Met Tyr Ser Tyr  
 345 350 355  
 cag gac ttg ctg acc ttt agc cta cct caa gac ggc caa act tac act 1158  
 Gln Asp Leu Leu Thr Phe Ser Leu Pro Gln Asp Gly Gln Thr Tyr Thr  
 360 365 370  
 aca gat aaa agt cac cag gtc cca ggc cgt ttt gtc tta gac cgg gtg 1206  
 Thr Asp Lys Ser His Gln Val Pro Gly Arg Phe Val Leu Asp Arg Val  
 375 380 385  
 gcc gac cat atc cgt tac tac cgg cgc ttg gct taa 1242  
 Ala Asp His Ile Arg Tyr Tyr Arg Arg Leu Ala  
 390 395 400

&lt;210&gt; 36

&lt;211&gt; 400

&lt;212&gt; PRT

<213> *Alloiococcus otitidis*

&lt;400&gt; 36

Met Val Ser Asp Tyr Phe Arg Arg Phe Asn Met Gln Ile Gly Ile Asp  
 1 5 10 15

Lys Leu Ala Phe Ala Thr Pro Thr Arg Tyr Leu Glu Met Ala Ser Leu  
 20 25 30

Ala Gln Ala Arg Ser Gln Asp Pro Asn Lys Tyr Ile Lys Gly Leu Gly  
 35 40 45

Gln Glu Ala Met Ala Val Pro Glu Glu Ser Asp Asp Ala Val Ser Leu  
 50 55 60

Ala Ala Asn Ala Gly Asn Leu Ile Leu Ser Glu Glu Asp Lys Ala Ala  
 65 70 75 80

Ile Asp Met Val Ile Val Gly Thr Glu Ser Gly Val Asp Gln Ser Lys  
85 90 95

Ser Ala Ala Ser Trp Val His Asp Leu Leu Gly Ile Asn Pro His Ala  
100 105 110

Arg Ser Leu Glu Ile Lys Gln Ala Cys Tyr Gly Ala Thr Ala Gly Leu  
115 120 125

Lys Leu Ala Val Ala His Leu Ala Leu Asn Pro Asp Ser Lys Val Leu  
130 135 140

Val Ile Gly Ser Asp Ile Ala Lys Tyr Gly Leu Glu Thr Gly Gly Glu  
145 150 155 160

Pro Thr Gln Gly Ala Gly Ala Val Ala Ile Leu Val Ser Arg Asp Pro  
165 170 175

Ala Ile Ala Val Val Asn Asn Asp Ser Ala Met Leu Thr Lys Asn Ile  
180 185 190

Ala Asp Phe Trp Arg Pro Asn Tyr Ser Asp Tyr Ala His Val Asp Gly  
195 200 205

Lys Phe Ser Asn Gln Ala Tyr Leu Ser Asn Leu Ala Glu Val Trp Arg  
210 215 220

Gln Tyr Lys Ile Lys Asn Gln Leu Ser Ala Lys Asp Phe Lys Ala Met  
225 230 235 240

Val Phe His Ser Pro Tyr Thr Lys Met Gly Lys Lys Ala Leu Leu Lys  
245 250 255

Leu Gly Asp Tyr Glu Asp Gln Lys Glu Ile Asp Arg Leu Leu Ala Tyr  
260 265 270

Tyr Glu Pro Gly Arg Tyr Tyr Asn Lys Arg Val Gly Asn Ile Tyr Thr  
275 280 285

Gly Ser Leu Tyr Leu Ser Leu Ile Ser Leu Leu Asp Gln Val Ser Asp  
290 295 300

Leu Glu Ala Gly Asp Arg Ile Gly Leu Tyr Ser Tyr Gly Ser Gly Ala  
 305 310 315 320

Val Gly Glu Phe Phe Ser Ile Arg Leu Gln Pro Gly Tyr Lys Glu Ser  
 325 330 335

Leu Gln Gln Val Asp Phe Asp Gln Val Val Asn Gln Arg Ser Ala Leu  
 340 345 350

Glu Met Tyr Ser Tyr Gln Asp Leu Leu Thr Phe Ser Leu Pro Gln Asp  
 355 360 365

Gly Gln Thr Tyr Thr Thr Asp Lys Ser His Gln Val Pro Gly Arg Phe  
 370 375 380

Val Leu Asp Arg Val Ala Asp His Ile Arg Tyr Tyr Arg Arg Leu Ala  
 385 390 395 400

<210> 37

<211> 1323

<212> DNA

<213> *Alloioococcus otitidis*

<220>

<221> CDS

<222> (31)..(1323)

<223>

<400> 37

ttctggtata gattaaggaa ggaggagacc atg tta ccc tta ttc aag caa ttt 54  
 Met Leu Pro Leu Phe Lys Gln Phe  
 1 5

tac aag caa agc ctc agc cag cgc ctc aaa gct cta gaa aag gcc ggc 102  
 Tyr Lys Gln Ser Leu Ser Gln Arg Leu Lys Ala Leu Glu Lys Ala Gly  
 10 15 20

tat ctt gat cct gac cag gcg ggt aaa ctc cag tca ggg gaa ctg ggt 150  
 Tyr Leu Asp Pro Asp Gln Ala Gly Lys Leu Gln Ser Gly Glu Leu Gly  
 25 30 35 40

ttg acc cat gaa gcc ggc gac cac atg att gaa aac tac atc ggc tcc 198  
 Leu Thr His Glu Ala Gly Asp His Met Ile Glu Asn Tyr Ile Gly Ser  
 45 50 55

tat acc ctc cct ctg gga ctg gcc ctc cac ttt tta ctc gat ggc aag 246  
 Tyr Thr Leu Pro Leu Gly Leu Ala Leu His Phe Leu Leu Asp Gly Lys

60	65	70	
agc tac cta gtc ccc atg gct att gaa gag ccc tct gtc att gcc gct			294
Ser Tyr Leu Val Pro Met Ala Ile Glu Glu Pro Ser Val Ile Ala Ala			
75	80	85	
gcc agc aac ggt gcc aag atg gta gcc caa agc ggt ggt ttc cat aca			342
Ala Ser Asn Gly Ala Lys Met Val Ala Gln Ser Gly Gly Phe His Thr			
90	95	100	
gtc aag gaa aac cgg ctg atg atc ggt caa gtg gtc ata gcc gga agc			390
Val Lys Glu Asn Arg Leu Met Ile Gly Gln Val Val Ile Ala Gly Ser			
105	110	115	120
aca aaa cct agc cag gac cgg gga aaa atc ctg agc cac cag caa gac			438
Thr Lys Pro Ser Gln Asp Arg Gly Lys Ile Leu Ser His Gln Gln Asp			
125	130	135	
tta atc gac cta gcc aat gct agc tat ccc tca att ggt aaa aga ggg			486
Leu Ile Asp Leu Ala Asn Ala Ser Tyr Pro Ser Ile Gly Lys Arg Gly			
140	145	150	
ggg ggg gcc cga ggc att caa gtc aaa cag ttt gac tca gac ctg ggc			534
Gly Gly Ala Arg Gly Ile Gln Val Lys Gln Phe Asp Ser Asp Leu Gly			
155	160	165	
cag gat atg gga agc tat ctg gca gtc tac ttg act gtt gac tgc cag			582
Gln Asp Met Gly Ser Tyr Leu Ala Val Tyr Leu Thr Val Asp Cys Gln			
170	175	180	
gaa gcc atg ggg gct aac att atc aac acc atg ctg gaa gcc ctg gct			630
Glu Ala Met Gly Ala Asn Ile Ile Asn Thr Met Leu Glu Ala Leu Ala			
185	190	195	200
cct gaa att gac cgc cta acc agc ggc cag gtc ttg atg tcc atc tta			678
Pro Glu Ile Asp Arg Leu Thr Ser Gly Gln Val Leu Met Ser Ile Leu			
205	210	215	
tct aac ctg gcc act gaa tcc ctt gtc act gtt tcc tgt caa gta aaa			726
Ser Asn Leu Ala Thr Glu Ser Leu Val Thr Val Ser Cys Gln Val Lys			
220	225	230	
ccc aga ttt tta gtc aaa aat gac atg gca ggg gaa gct gtc cgg gac			774
Pro Arg Phe Leu Val Lys Asn Asp Met Ala Gly Glu Ala Val Arg Asp			
235	240	245	
caa atc atc cag gcc tac cag tat gcc tgc ctg gac ccc tac cgg gca			822
Gln Ile Ile Gln Ala Tyr Gln Tyr Ala Cys Leu Asp Pro Tyr Arg Ala			
250	255	260	
gcc acc cac aac aag ggg atc atg aac ggg gta gac ggc ttg gtc cta			870
Ala Thr His Asn Lys Gly Ile Met Asn Gly Val Asp Gly Leu Val Leu			
265	270	275	280
gct agt ggg aat gat tgg cgg gca atc gaa gcg ggg gcc cat gct tac			918
Ala Ser Gly Asn Asp Trp Arg Ala Ile Glu Ala Gly Ala His Ala Tyr			
285	290	295	

gct agt ttg acc ggc cac tac cgc ccc ttg tcc aag tgg gaa aag acc 966  
 Ala Ser Leu Thr Gly His Tyr Arg Pro Leu Ser Lys Trp Glu Lys Thr  
 300 305 310  
 caa gac gga cag tta aaa ggg acc att acc ctt ccc ttg cca att gcc 1014  
 Gln Asp Gly Gln Leu Lys Gly Thr Ile Thr Leu Pro Leu Pro Ile Ala  
 315 320 325  
 aca gtt ggt ggg gct att gcc tcc cac cct gta gcc caa gtt agc cag 1062  
 Thr Val Gly Gly Ala Ile Ala Ser His Pro Val Ala Gln Val Ser Gln  
 330 335 340  
 caa atc tta ggc caa cct act gct aag caa tta gcc cgg ctg gtt gca 1110  
 Gln Ile Leu Gly Gln Pro Thr Ala Lys Gln Leu Ala Arg Leu Val Ala  
 345 350 355 360  
 gca gtg gga cta gcc cag aac cta tcc gct ctt cgt gcc tta gtc aca 1158  
 Ala Val Gly Leu Ala Gln Asn Leu Ser Ala Leu Arg Ala Leu Val Thr  
 365 370 375  
 act ggt att caa caa gga cac atg gcc ctc cag gca agg tct ttg gcc 1206  
 Thr Gly Ile Gln Gln Gly His Met Ala Leu Gln Ala Arg Ser Leu Ala  
 380 385 390  
 atg aat gcc ggg gcc cgg gga gac aag atc caa aag ctg gca gac cgc 1254  
 Met Asn Ala Gly Ala Arg Gly Asp Lys Ile Gln Lys Leu Ala Asp Arg  
 395 400 405  
 tta att aac caa gac caa atg aac cta gca act gcc cgt gcc ctg ctc 1302  
 Leu Ile Asn Gln Asp Gln Met Asn Leu Ala Thr Ala Arg Ala Leu Leu  
 410 415 420  
 aag aac atg gaa gaa gac taa 1323  
 Lys Asn Met Glu Glu Asp  
 425 430

&lt;210&gt; 38

&lt;211&gt; 430

&lt;212&gt; PRT

<213> *Alloiococcus otitidis*

&lt;400&gt; 38

Met Leu Pro Leu Phe Lys Gln Phe Tyr Lys Gln Ser Leu Ser Gln Arg  
 1 5 10 15

Leu Lys Ala Leu Glu Lys Ala Gly Tyr Leu Asp Pro Asp Gln Ala Gly  
 20 25 30

Lys Leu Gln Ser Gly Glu Leu Gly Leu Thr His Glu Ala Gly Asp His  
 35 40 45

Met Ile Glu Asn Tyr Ile Gly Ser Tyr Thr Leu Pro Leu Gly Leu Ala

50

55

60

Leu His Phe Leu Leu Asp Gly Lys Ser Tyr Leu Val Pro Met Ala Ile  
65 70 75 80

Glu Glu Pro Ser Val Ile Ala Ala Ala Ser Asn Gly Ala Lys Met Val  
85 90 95

Ala Gln Ser Gly Gly Phe His Thr Val Lys Glu Asn Arg Leu Met Ile  
100 105 110

Gly Gln Val Val Ile Ala Gly Ser Thr Lys Pro Ser Gln Asp Arg Gly  
115 120 125

Lys Ile Leu Ser His Gln Gln Asp Leu Ile Asp Leu Ala Asn Ala Ser  
130 135 140

Tyr Pro Ser Ile Gly Lys Arg Gly Gly Gly Ala Arg Gly Ile Gln Val  
145 150 155 160

Lys Gln Phe Asp Ser Asp Leu Gly Gln Asp Met Gly Ser Tyr Leu Ala  
165 170 175

Val Tyr Leu Thr Val Asp Cys Gln Glu Ala Met Gly Ala Asn Ile Ile  
180 185 190

Asn Thr Met Leu Glu Ala Leu Ala Pro Glu Ile Asp Arg Leu Thr Ser  
195 200 205

Gly Gln Val Leu Met Ser Ile Leu Ser Asn Leu Ala Thr Glu Ser Leu  
210 215 220

Val Thr Val Ser Cys Gln Val Lys Pro Arg Phe Leu Val Lys Asn Asp  
225 230 235 240

Met Ala Gly Glu Ala Val Arg Asp Gln Ile Ile Gln Ala Tyr Gln Tyr  
245 250 255

Ala Cys Leu Asp Pro Tyr Arg Ala Ala Thr His Asn Lys Gly Ile Met  
260 265 270

Asn Gly Val Asp Gly Leu Val Leu Ala Ser Gly Asn Asp Trp Arg Ala  
275 280 285

Ile Glu Ala Gly Ala His Ala Tyr Ala Ser Leu Thr Gly His Tyr Arg  
 290 295 300

Pro Leu Ser Lys Trp Glu Lys Thr Gln Asp Gly Gln Leu Lys Gly Thr  
 305 310 315 320

Ile Thr Leu Pro Leu Pro Ile Ala Thr Val Gly Gly Ala Ile Ala Ser  
 325 330 335

His Pro Val Ala Gln Val Ser Gln Gln Ile Leu Gly Gln Pro Thr Ala  
 340 345 350

Lys Gln Leu Ala Arg Leu Val Ala Ala Val Gly Leu Ala Gln Asn Leu  
 355 360 365

Ser Ala Leu Arg Ala Leu Val Thr Thr Gly Ile Gln Gln Gly His Met  
 370 375 380

Ala Leu Gln Ala Arg Ser Leu Ala Met Asn Ala Gly Ala Arg Gly Asp  
 385 390 395 400

Lys Ile Gln Lys Leu Ala Asp Arg Leu Ile Asn Gln Asp Gln Met Asn  
 405 410 415

Leu Ala Thr Ala Arg Ala Leu Leu Lys Asn Met Glu Glu Asp  
 420 425 430

<210> 39

<211> 930

<212> DNA

<213> *Alloiococcus otitidis*

<220>

<221> CDS

<222> (13)..(930)

<223>

<400> 39

aggattagta aa atg tta ttt gat cgt atc gta gaa gcc ttt ccc gaa agc 51  
 Met Leu Phe Asp Arg Ile Val Glu Ala Phe Pro Glu Ser  
 1 5 10

aac atc aaa aaa gat gaa ccc ttg tcc tat tac tct tac act cga aca 99  
 Asn Ile Lys Lys Asp Glu Pro Leu Ser Tyr Tyr Ser Tyr Thr Arg Thr  
 15 20 25



ggt ggc ccg gct gac att ttg att ttc cca gaa tcc atc gat gaa att Gly Gly Pro Ala Asp Ile Leu Ile Phe Pro Glu Ser Ile Asp Glu Ile 30 35 40 45	147
gtg acg att atc aag tgg atc aac caa agt ccg gaa tac caa gct ggc Val Thr Ile Ile Lys Trp Ile Asn Gln Ser Pro Glu Tyr Gln Ala Gly 50 55 60	195
gat ctc ccc ctc act atc tta ggc aat gct agc aac ctg atc gta aaa Asp Leu Pro Leu Thr Ile Leu Gly Asn Ala Ser Asn Leu Ile Val Lys 65 70 75	243
gat ggt ggg ata aga ggg att acc atc att acc acc ggc att aaa acc Asp Gly Gly Ile Arg Gly Ile Thr Ile Ile Thr Thr Gly Ile Lys Thr 80 85 90	291
att tgt cac gaa gag aac cgg atc act gcg ggc gct gga gca gct att Ile Cys His Glu Glu Asn Arg Ile Thr Ala Gly Ala Gly Ala Ala Ile 95 100 105	339
atc gat gtt agc cag gct gcc ttg gac cat agc tta act ggc ttg gaa Ile Asp Val Ser Gln Ala Ala Leu Asp His Ser Leu Thr Gly Leu Glu 110 115 120 125	387
ttc gct tgt ggc ata ccg ggt agt aca ggc ggg gct gtt tac atg aac Phe Ala Cys Gly Ile Pro Gly Ser Thr Gly Gly Ala Val Tyr Met Asn 130 135 140	435
gct ggg gct tac ggt ggg gaa gtc cag cat tgt gtt gaa agt gtc caa Ala Gly Ala Tyr Gly Gly Glu Val Gln His Cys Val Glu Ser Val Gln 145 150 155	483
gtc ttg acc cgg cat ggc cag ttg aag acc tat agt aat gcg gaa atg Val Leu Thr Arg His Gly Gln Leu Lys Thr Tyr Ser Asn Ala Glu Met 160 165 170	531
aac ttc tcc tac cgc cac agt tat ttg atg gaa gaa gac gat ata gta Asn Phe Ser Tyr Arg His Ser Tyr Leu Met Glu Glu Asp Asp Ile Val 175 180 185	579
gtc tcc gtg acc ttt aaa ttg gag tcg ggc gac tac atc act atc aag Val Ser Val Thr Phe Lys Leu Glu Ser Gly Asp Tyr Ile Thr Ile Lys 190 195 200 205	627
gaa aag atg gat gaa tta acc tac ctt aga gaa tcc aaa caa ccg ctg Glu Lys Met Asp Glu Leu Thr Tyr Leu Arg Glu Ser Lys Gln Pro Leu 210 215 220	675
gaa tac ccc tct tgt ggg tca gtc ttt aaa aga cct gaa ggc cac ttt Glu Tyr Pro Ser Cys Gly Ser Val Phe Lys Arg Pro Glu Gly His Phe 225 230 235	723
acg ggg aaa tta atc cag gat gct ggc ctt caa gga ttg gtc cat ggt Thr Gly Lys Leu Ile Gln Asp Ala Gly Leu Gln Gly Leu Val His Gly 240 245 250	771
gga gcc cag gta tcc gaa aaa cat gcc ggt ttt atc att aat ata ggc	819

Gly Ala Gln Val Ser Glu Lys His Ala Gly Phe Ile Ile Asn Ile Gly  
 255 260 265

aat gct acc gcc agc gac tac caa gag ttg atc caa cat atc caa gaa 867  
 Asn Ala Thr Ala Ser Asp Tyr Gln Glu Leu Ile Gln His Ile Gln Glu  
 270 275 280 285

gaa gtc tac cgg att tac aag gtt aag ctg gaa cgt gaa gtt cgc att 915  
 Glu Val Tyr Arg Ile Tyr Lys Val Lys Leu Glu Arg Glu Val Arg Ile  
 290 295 300

ata ggg gag gat tag 930  
 Ile Gly Glu Asp  
 305

<210> 40

<211> 305

<212> PRT

<213> *Alloiococcus otitidis*

<400> 40

Met Leu Phe Asp Arg Ile Val Glu Ala Phe Pro Glu Ser Asn Ile Lys  
 1 5 10 15

Lys Asp Glu Pro Leu Ser Tyr Tyr Ser Tyr Thr Arg Thr Gly Gly Pro  
 20 25 30

Ala Asp Ile Leu Ile Phe Pro Glu Ser Ile Asp Glu Ile Val Thr Ile  
 35 40 45

Ile Lys Trp Ile Asn Gln Ser Pro Glu Tyr Gln Ala Gly Asp Leu Pro  
 50 55 60

Leu Thr Ile Leu Gly Asn Ala Ser Asn Leu Ile Val Lys Asp Gly Gly  
 65 70 75 80

Ile Arg Gly Ile Thr Ile Ile Thr Thr Gly Ile Lys Thr Ile Cys His  
 85 90 95

Glu Glu Asn Arg Ile Thr Ala Gly Ala Gly Ala Ala Ile Ile Asp Val  
 100 105 110

Ser Gln Ala Ala Leu Asp His Ser Leu Thr Gly Leu Glu Phe Ala Cys  
 115 120 125

Gly Ile Pro Gly Ser Thr Gly Gly Ala Val Tyr Met Asn Ala Gly Ala  
 130 135 140

Tyr Gly Gly Glu Val Gln His Cys Val Glu Ser Val Gln Val Leu Thr  
 145 150 155 160

Arg His Gly Gln Leu Lys Thr Tyr Ser Asn Ala Glu Met Asn Phe Ser  
 165 170 175

Tyr Arg His Ser Tyr Leu Met Glu Glu Asp Asp Ile Val Val Ser Val  
 180 185 190

Thr Phe Lys Leu Glu Ser Gly Asp Tyr Ile Thr Ile Lys Glu Lys Met  
 195 200 205

Asp Glu Leu Thr Tyr Leu Arg Glu Ser Lys Gln Pro Leu Glu Tyr Pro  
 210 215 220

Ser Cys Gly Ser Val Phe Lys Arg Pro Glu Gly His Phe Thr Gly Lys  
 225 230 235 240

Leu Ile Gln Asp Ala Gly Leu Gln Gly Leu Val His Gly Gly Ala Gln  
 245 250 255

Val Ser Glu Lys His Ala Gly Phe Ile Ile Asn Ile Gly Asn Ala Thr  
 260 265 270

Ala Ser Asp Tyr Gln Glu Leu Ile Gln His Ile Gln Glu Glu Val Tyr  
 275 280 285

Arg Ile Tyr Lys Val Lys Leu Glu Arg Glu Val Arg Ile Ile Gly Glu  
 290 295 300

Asp  
 305

<210> 41  
 <211> 1104  
 <212> DNA  
 <213> *Alloiococcus otitidis*

<220>  
 <221> CDS  
 <222> (16) .. (1104)  
 <223>

<400> 41  
 aaagctggtg ttttc atg gtt tat agc tta agg att ccg ggg aaa ctt tat

Met Val Tyr Ser Leu Arg Ile Pro Gly Lys Leu Tyr															
1 5 10															
ttg gca ggt gaa tac gca gta gta acc ccc ggc tat gcc ggg atc ttg	99														
Leu Ala Gly Glu Tyr Ala Val Val Thr Pro Gly Tyr Ala Gly Ile Leu															
15 20 25															
ctg aca gtc agc cgg tat ttg act tta gac att tgg gaa aca tct ccc	147														
Leu Thr Val Ser Arg Tyr Leu Thr Leu Asp Ile Trp Glu Thr Ser Pro															
30 35 40															
gac caa gct tca gtc agg tct caa aca tat ggc aac cag gcc tat gct	195														
Asp Gln Ala Ser Val Arg Ser Gln Thr Tyr Gly Asn Gln Ala Tyr Ala															
45 50 55 60															
tgg gag cgg tta gat ggt atc ttt agc ttt aag gac tgg tcc cac ccc	243														
Trp Glu Arg Leu Asp Gly Ile Phe Ser Phe Lys Asp Trp Ser His Pro															
65 70 75															
ttc cac cta gtc gaa acg gtg atc caa aca gtg gaa gcc tac ata gaa	291														
Phe His Leu Val Glu Thr Val Ile Gln Thr Val Glu Ala Tyr Ile Glu															
80 85 90															
tcc ttg tcc ctg cct tta aaa agt tac ggg att cag atc aag agc cag	339														
Ser Leu Ser Leu Pro Leu Lys Ser Tyr Gly Ile Gln Ile Lys Ser Gln															
95 100 105															
ttg gac tac cag ggc aaa aaa att ggc ctg ggg tct agt ggg gcc gtt	387														
Leu Asp Tyr Gln Gly Lys Lys Ile Gly Leu Gly Ser Ser Gly Ala Val															
110 115 120															
acc atc gct gtt atc cga ggc ctg agc ctt ctt tac gac ctc cac tta	435														
Thr Ile Ala Val Ile Arg Gly Leu Ser Leu Leu Tyr Asp Leu His Leu															
125 130 135 140															
aaa gac ata gac att ttc aaa cta gct gcc atc gcc cat atc cag cta	483														
Lys Asp Ile Asp Ile Phe Lys Leu Ala Ala Ile Ala His Ile Gln Leu															
145 150 155															
aag agc aag ggg tct ttt ggc gat ttg gca gcc tgc act tat act ggt	531														
Lys Ser Lys Gly Ser Phe Gly Asp Leu Ala Ala Cys Thr Tyr Thr Gly															
160 165 170															
gtg atc cgc tac cag tcc ctg gat aga gaa tgg tta caa gaa caa atc	579														
Val Ile Arg Tyr Gln Ser Leu Asp Arg Glu Trp Leu Gln Glu Gln Ile															
175 180 185															
tcc aac cat tcc atc aag gac ctc ctg gcc atg gat tgg cct agc cta	627														
Ser Asn His Ser Ile Lys Asp Leu Leu Ala Met Asp Trp Pro Ser Leu															
190 195 200															
ggt cta gac cgg ctc agc ctg ccc cat gac ctc agg ctt tta atc gga	675														
Gly Leu Asp Arg Leu Ser Leu Pro His Asp Leu Arg Leu Leu Ile Gly															
205 210 215 220															
tgg acc ggc cag cct gcc tcc aca gaa aaa ttg gtt cag gct gtc tac	723														
Trp Thr Gly Gln Pro Ala Ser Thr Glu Lys Leu Val Gln Ala Val Tyr															

Val Arg Ser Gln Thr Tyr Gly Asn Gln Ala Tyr Ala Trp Glu Arg Leu  
50 55 60

Asp Gly Ile Phe Ser Phe Lys Asp Trp Ser His Pro Phe His Leu Val  
65 70 75 80

Glu Thr Val Ile Gln Thr Val Glu Ala Tyr Ile Glu Ser Leu Ser Leu  
85 90 95

Pro Leu Lys Ser Tyr Gly Ile Gln Ile Lys Ser Gln Leu Asp Tyr Gln  
100 105 110

Gly Lys Lys Ile Gly Leu Gly Ser Ser Gly Ala Val Thr Ile Ala Val  
115 120 125

Ile Arg Gly Leu Ser Leu Leu Tyr Asp Leu His Leu Lys Asp Ile Asp  
130 135 140

Ile Phe Lys Leu Ala Ala Ile Ala His Ile Gln Leu Lys Ser Lys Gly  
145 150 155 160

Ser Phe Gly Asp Leu Ala Ala Cys Thr Tyr Thr Gly Val Ile Arg Tyr  
165 170 175

Gln Ser Leu Asp Arg Glu Trp Leu Gln Glu Gln Ile Ser Asn His Ser  
180 185 190

Ile Lys Asp Leu Leu Ala Met Asp Trp Pro Ser Leu Gly Leu Asp Arg  
195 200 205

Leu Ser Leu Pro His Asp Leu Arg Leu Leu Ile Gly Trp Thr Gly Gln  
210 215 220

Pro Ala Ser Thr Glu Lys Leu Val Gln Ala Val Tyr Pro Gln Lys Ile  
225 230 235 240

Thr Arg Thr Pro Leu Asp Phe Gln Ser Phe Leu Asp Gln Ser Gln Glu  
245 250 255

Cys Val Asp Gly Leu Val Glu Ser Leu Ser Gln Ala Asp Ser Gln Ala  
260 265 270

Ser Leu Ala Trp Ile Gln Lys Asn Arg Thr Leu Leu Lys Ala Met Gly  
275 280 285

Gln Ser Arg Gly Lys Val Ile Glu Thr Lys Ala Leu Thr Tyr Leu Cys

290

295

300

Asp Ile Val Ala Lys Tyr Gly Gly Gln Ala Lys Ser Ser Gly Ala Gly  
 305 310 315 320

Gly Gly Asp Cys Gly Ile Gly Leu Ile Thr Arg Glu Ser Pro Ile Glu  
 325 330 335

Ala Ile Tyr Arg Glu Trp Met Asp Ala Gly Ile Leu Pro Leu Arg Leu  
 340 345 350

Asp Ile Val Glu Asn Gly Ala Cys Tyr Asp  
 355 360

&lt;210&gt; 43

&lt;211&gt; 1023

&lt;212&gt; DNA

<213> *Alloiococcus otitidis*

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (13)..(1023)

&lt;223&gt;

&lt;400&gt; 43

gagaagccaa cc atg act aag cag gcc ttt gaa aag aaa aag tta ggc cgg 51  
 Met Thr Lys Gln Ala Phe Glu Lys Lys Lys Leu Gly Arg  
 1 5 10

att tgc cgg gcc cat acc aac att gcc ttg atc aag tac tgg ggt aag 99  
 Ile Cys Arg Ala His Thr Asn Ile Ala Leu Ile Lys Tyr Trp Gly Lys  
 15 20 25

gct gat agg gac ttg att atc ccc aat aac aac tcc cta tct tta acc 147  
 Ala Asp Arg Asp Leu Ile Ile Pro Asn Asn Asn Ser Leu Ser Leu Thr  
 30 35 40 45

ttg gac gct ttt tat acc gat acc cag gta gtt ttt gac cca gac ttg 195  
 Leu Asp Ala Phe Tyr Thr Asp Thr Gln Val Val Phe Asp Pro Asp Leu  
 50 55 60

gac cag gac caa tta tgg cta gac ggg aaa cag gaa aaa ggg tcc gcc 243  
 Asp Gln Asp Gln Leu Trp Leu Asp Gly Lys Gln Glu Lys Gly Ser Ala  
 65 70 75

tta acc aag gcc cag gtc atc ctg gac ttg gtt cgg gac caa gcc cag 291  
 Leu Thr Lys Ala Gln Val Ile Leu Asp Leu Val Arg Asp Gln Ala Gln  
 80 85 90

ctt gac tgg ccg gcc aaa att acc agc cac aac caa gtt gcc act gca 339  
 Leu Asp Trp Pro Ala Lys Ile Thr Ser His Asn Gln Val Ala Thr Ala  
 95 100 105

gct ggc ttg gct tcc tct gct tct ggt ctg gcc gcc ttg gcg ggt gct Ala Gly Leu Ala Ser Ser Ala Ser Gly Leu Ala Ala Leu Ala Gly Ala 110 115 120 125	387
tca gct gat gct tta gac ctt ggc cta tcc cca act gac ctc tcc cga Ser Ala Asp Arg Ala Leu Asp Leu Gly Leu Ser Pro Thr Asp Leu Ser Arg 130 135 140	435
ttg gcc cgc agg gga tct ggg tct gcc tca cga agt att ttt ggt ggt Leu Ala Arg Arg Gly Ser Gly Ser Ala Ser Arg Ser Ile Phe Gly Gly 145 150 155	483
ttt gtc gag tgg gaa aag ggt cat gat gat agc tct tcc ttt gcc aag Phe Val Glu Trp Glu Lys Gly His Asp Asp Ser Ser Phe Ala Lys 160 165 170	531
ccc atc gac ttg gcc cag tgg gat att gcc atg ctc ttt gtc att gta Pro Ile Asp Leu Ala Gln Trp Asp Ile Ala Met Leu Phe Val Ile Val 175 180 185	579
agc gac cga cca aag gca att tcc tcc agc caa ggc atg caa ttg acc Ser Asp Arg Pro Lys Ala Ile Ser Ser Ser Gln Gly Met Gln Leu Thr 190 195 200 205	627
cag gag acg tcg gac ttt tac cag gcc tgg tta gac agc ctg gac caa Gln Glu Thr Ser Asp Phe Tyr Gln Ala Trp Leu Asp Ser Leu Asp Gln 210 215 220	675
gac cta gca gac atc aag tcc gct atc caa gcc caa gac ctc gac cag Asp Leu Ala Asp Ile Lys Ser Ala Ile Gln Ala Gln Asp Leu Asp Gln 225 230 235	723
gtt ggg tcc att gca gaa aga aat gcc ctg aaa atg cat gcc acc aac Val Gly Ser Ile Ala Glu Arg Asn Ala Leu Lys Met His Ala Thr Asn 240 245 250	771
ctg gca gcc aag ccc ccc ttc acc tat tgg act aaa gaa agt tta gcc Leu Ala Ala Lys Pro Pro Phe Thr Tyr Trp Thr Lys Glu Ser Leu Ala 255 260 265	819
ctg atg cag gaa gta tgg gac cgg cgc aag gct ggc cag tcc ctc tac Leu Met Gln Glu Val Trp Asp Arg Arg Lys Ala Gly Gln Ser Leu Tyr 270 275 280 285	867
ttc acc atg gac gcc ggc ccc aat gtc aag gtt att ggc agg gaa gct Phe Thr Met Asp Ala Gly Pro Asn Val Lys Val Ile Gly Arg Glu Ala 290 295 300	915
gac ctt aaa gcc ttc aaa gca gac ctc agc caa gac tgg ccc gac aag Asp Leu Lys Ala Phe Lys Ala Asp Leu Ser Gln Asp Trp Pro Asp Lys 305 310 315	963
cat ctt gtc tta gct aaa ccg ggt cca ggc ctg gcc ttt att gat gga His Leu Val Leu Ala Lys Pro Gly Pro Gly Leu Ala Phe Ile Asp Gly 320 325 330	1011



1023

cct ttg aac tag  
Pro Leu Asn  
335

<210> 44  
<211> 336  
<212> PRT  
<213> *Alloiococcus otitidis*

<400> 44  
Met Thr Lys Gln Ala Phe Glu Lys Lys Lys Leu Gly Arg Ile Cys Arg  
1 5 10 15

Ala His Thr Asn Ile Ala Leu Ile Lys Tyr Trp Gly Lys Ala Asp Arg  
20 25 30

Asp Leu Ile Ile Pro Asn Asn Asn Ser Leu Ser Leu Thr Leu Asp Ala  
35 40 45

Phe Tyr Thr Asp Thr Gln Val Val Phe Asp Pro Asp Leu Asp Gln Asp  
50 55 60

Gln Leu Trp Leu Asp Gly Lys Gln Glu Lys Gly Ser Ala Leu Thr Lys  
65 70 75 80

Ala Gln Val Ile Leu Asp Leu Val Arg Asp Gln Ala Gln Leu Asp Trp  
85 90 95

Pro Ala Lys Ile Thr Ser His Asn Gln Val Ala Thr Ala Ala Gly Leu  
100 105 110

Ala Ser Ser Ala Ser Gly Leu Ala Ala Leu Ala Gly Ala Ser Ala Asp  
115 120 125

Ala Leu Asp Leu Gly Leu Ser Pro Thr Asp Leu Ser Arg Leu Ala Arg  
130 135 140

Arg Gly Ser Gly Ser Ala Ser Arg Ser Ile Phe Gly Gly Phe Val Glu  
145 150 155 160

Trp Glu Lys Gly His Asp Asp Ser Ser Ser Phe Ala Lys Pro Ile Asp  
165 170 175

Leu Ala Gln Trp Asp Ile Ala Met Leu Phe Val Ile Val Ser Asp Arg  
180 185 190

Pro Lys Ala Ile Ser Ser Ser Gln Gly Met Gln Leu Thr Gln Glu Thr  
 195 200 205

Ser Asp Phe Tyr Gln Ala Trp Leu Asp Ser Leu Asp Gln Asp Leu Ala  
 210 215 220

Asp Ile Lys Ser Ala Ile Gln Ala Gln Asp Leu Asp Gln Val Gly Ser  
 225 230 235 240

Ile Ala Glu Arg Asn Ala Leu Lys Met His Ala Thr Asn Leu Ala Ala  
 245 250 255

Lys Pro Pro Phe Thr Tyr Trp Thr Lys Glu Ser Leu Ala Leu Met Gln  
 260 265 270

Glu Val Trp Asp Arg Arg Lys Ala Gly Gln Ser Leu Tyr Phe Thr Met  
 275 280 285

Asp Ala Gly Pro Asn Val Lys Val Ile Gly Arg Glu Ala Asp Leu Lys  
 290 295 300

Ala Phe Lys Ala Asp Leu Ser Gln Asp Trp Pro Asp Lys His Leu Val  
 305 310 315 320

Leu Ala Lys Pro Gly Pro Gly Leu Ala Phe Ile Asp Gly Pro Leu Asn  
 325 330 335

<210> 45

<211> 981

<212> DNA

<213> *Alloiococcus otitidis*

<220>

<221> CDS

<222> (28)..(981)

<223>

<400> 45

acaaaaatag acaaaggaga caaaagg atg acg ctt gtt aaa aat gta gcc aaa  
 Met Thr Leu Val Lys Asn Val Ala Lys  
 1 5

54

ggc act gcc cat ggt aaa att att tta atc ggt gag cat gct gtt gtc

102

Gly Thr Ala His Gly Lys Ile Ile Leu Ile Gly Glu His Ala Val Val	
10 15 20 25	
tat aac atg ccg gcc atc gcc ctc cct ttt acc aca gcc acc atc acc	150
Tyr Asn Met Pro Ala Ile Ala Leu Pro Phe Thr Thr Ala Thr Ile Thr	
30 35 40	
gtt gaa gtt agt cct tac caa ggc aaa agc tat cta gaa agt gct tgc	198
Val Glu Val Ser Pro Tyr Gln Gly Lys Ser Tyr Leu Glu Ser Ala Cys	
45 50 55	
tac tgc gga tct tta gac caa gcg ccc ggg gac ttg gca ggg ctt caa	246
Tyr Cys Gly Ser Leu Asp Gln Ala Pro Gly Asp Leu Ala Gly Leu Gln	
60 65 70	
gcc tgt ttg aca gcg gtt tgt gcc gac tta gac cag tcc agc gac cac	294
Ala Cys Leu Thr Ala Val Cys Ala Asp Leu Asp Gln Ser Ser Asp His	
75 80 85	
ttg tat atc aag gtc gac agc atg atc cct gct gaa aga gga atg ggg	342
Leu Tyr Ile Lys Val Asp Ser Met Ile Pro Ala Glu Arg Gly Met Gly	
90 95 100 105	
tcc agt gct gct gtg gcc acc gcc tta gtc aag gcc ctc ttt cac tac	390
Ser Ser Ala Ala Val Ala Thr Ala Leu Val Lys Ala Leu Phe His Tyr	
110 115 120	
ttc caa gtc gac tta agc agt gaa gcc ctc tca gcc tat gtc gag att	438
Phe Gln Val Asp Leu Ser Ser Glu Ala Leu Ser Ala Tyr Val Glu Ile	
125 130 135	
gcc gaa aaa att acc cat ggc aag cca tcg ggt ctg gat gct aca gtc	486
Ala Glu Lys Ile Thr His Gly Lys Pro Ser Gly Leu Asp Ala Thr Val	
140 145 150	
gtc aac tcc att gcc ccc gtt tat ttt aaa cgc aac cag ctt ccc aag	534
Val Asn Ser Ile Ala Pro Val Tyr Phe Lys Arg Asn Gln Leu Pro Lys	
155 160 165	
gcc atc cct tta aat gtt gac ggc tat tta att gca gcc gat act ggg	582
Ala Ile Pro Leu Asn Val Asp Gly Tyr Leu Ile Ala Ala Asp Thr Gly	
170 175 180 185	
att aag ggc cac acg aaa gaa gcc gtt ggg gat gtg gcg aag ctg gtt	630
Ile Lys Gly His Thr Lys Glu Ala Val Gly Asp Val Ala Lys Leu Val	
190 195 200	
gaa act gcc aag gtt caa acc atg gac att gtc cac cac ctc ggc cag	678
Glu Thr Ala Lys Val Gln Thr Met Asp Ile Val His His Leu Gly Gln	
205 210 215	
ctt acc cac cag gct aaa aaa gca atc atg acc aat aac ctc cct ggc	726
Leu Thr His Gln Ala Lys Lys Ala Ile Met Thr Asn Asn Leu Pro Gly	
220 225 230	
tta ggg gag att ttg aac cag tcc cac caa ctc tta aag gat tta act	774
Leu Gly Glu Ile Leu Asn Gln Ser His Gln Leu Leu Lys Asp Leu Thr	

235 240 245

gtc agc aat ccc aag tta gac caa ctt gtc caa gca gcc caa gat gct 822  
Val Ser Asn Pro Lys Leu Asp Gln Leu Val Gln Ala Ala Gln Asp Ala  
250 255 260 265

gga gct tgc gga gct aag tta acc ggt ggg ggc cgg ggt ggt tgc atg 870  
Gly Ala Cys Gly Ala Lys Leu Thr Gly Gly Gly Arg Gly Gly Cys Met  
270 275 280

att gcc cta gcc caa agc aac cag gat gcc tcc aat att gcc caa aaa 918  
Ile Ala Leu Ala Gln Ser Asn Gln Asp Ala Ser Asn Ile Ala Gln Lys  
285 290 295

ttg gaa aaa gcg gga gcc att gaa acc tgg atc cac ccc tta gga gaa 966  
Leu Glu Lys Ala Gly Ala Ile Glu Thr Trp Ile His Pro Leu Gly Glu  
300 305 310

gcc aac cat gac taa 981  
Ala Asn His Asp  
315

<210> 46  
<211> 317  
<212> PRT  
<213> *Alloiococcus otitidis*

<400> 46  
Met Thr Leu Val Lys Asn Val Ala Lys Gly Thr Ala His Gly Lys Ile  
1 5 10 15

Ile Leu Ile Gly Glu His Ala Val Val Tyr Asn Met Pro Ala Ile Ala  
20 25 30

Leu Pro Phe Thr Thr Ala Thr Ile Thr Val Glu Val Ser Pro Tyr Gln  
35 40 45

Gly Lys Ser Tyr Leu Glu Ser Ala Cys Tyr Cys Gly Ser Leu Asp Gln  
50 55 60

Ala Pro Gly Asp Leu Ala Gly Leu Gln Ala Cys Leu Thr Ala Val Cys  
65 70 75 80

Ala Asp Leu Asp Gln Ser Ser Asp His Leu Tyr Ile Lys Val Asp Ser  
85 90 95

Met Ile Pro Ala Glu Arg Gly Met Gly Ser Ser Ala Ala Val Ala Thr  
100 105 110

Ala Leu Val Lys Ala Leu Phe His Tyr Phe Gln Val Asp Leu Ser Ser  
115 120 125

Glu Ala Leu Ser Ala Tyr Val Glu Ile Ala Glu Lys Ile Thr His Gly  
130 135 140

Lys Pro Ser Gly Leu Asp Ala Thr Val Val Asn Ser Ile Ala Pro Val  
145 150 155 160

Tyr Phe Lys Arg Asn Gln Leu Pro Lys Ala Ile Pro Leu Asn Val Asp  
165 170 175

Gly Tyr Leu Ile Ala Ala Asp Thr Gly Ile Lys Gly His Thr Lys Glu  
180 185 190

Ala Val Gly Asp Val Ala Lys Leu Val Glu Thr Ala Lys Val Gln Thr  
195 200 205

Met Asp Ile Val His His Leu Gly Gln Leu Thr His Gln Ala Lys Lys  
210 215 220

Ala Ile Met Thr Asn Asn Leu Pro Gly Leu Gly Glu Ile Leu Asn Gln  
225 230 235 240

Ser His Gln Leu Leu Lys Asp Leu Thr Val Ser Asn Pro Lys Leu Asp  
245 250 255

Gln Leu Val Gln Ala Ala Gln Asp Ala Gly Ala Cys Gly Ala Lys Leu  
260 265 270

Thr Gly Gly Gly Arg Gly Gly Cys Met Ile Ala Leu Ala Gln Ser Asn  
275 280 285

Gln Asp Ala Ser Asn Ile Ala Gln Lys Leu Glu Lys Ala Gly Ala Ile  
290 295 300

Glu Thr Trp Ile His Pro Leu Gly Glu Ala Asn His Asp  
305 310 315

<210> 47

<211> 975

<212> DNA

<213> Alloiococcus otitidis

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (46)..(975)

&lt;223&gt;

&lt;400&gt; 47

agaatcaaatt ttgtttttaa attatcagct tttggagggtc agaac atg aac aat tcc 57  
 Met Asn Asn Ser  
 1

cgt att ttt tta gtc tat gac cgt aaa gac tgg cag tct ctt aga gaa 105  
 Arg Ile Phe Leu Val Tyr Asp Arg Lys Asp Trp Gln Ser Leu Arg Glu  
 5 10 15 20

aat gcc agc ctt tct tta acg gaa aaa aac cta aat aac ttg cgt gca 153  
 Asn Ala Ser Leu Ser Leu Thr Glu Lys Asn Leu Asn Asn Leu Arg Ala  
 25 30 35

gtg aat gac gtc ata tcg atg gaa gat gtc cga gaa gtt tac gtc ccc 201  
 Val Asn Asp Val Ile Ser Met Glu Asp Val Arg Glu Val Tyr Val Pro  
 40 45 50

att atc caa tta ctg gat gtc tac ata aaa agt tac tac cgc cac cag 249  
 Ile Ile Gln Leu Leu Asp Val Tyr Ile Lys Ser Tyr Tyr Arg His Gln  
 55 60 65

gct tcc ttg atc aat tac ttg aac ctg gac cag cct aaa aag tac caa 297  
 Ala Ser Leu Ile Asn Tyr Leu Asn Leu Asp Gln Pro Lys Lys Tyr Gln  
 70 75 80

ccc tat gtg att ggg att gca ggg agc gtg gct gtg ggc aag tct acg 345  
 Pro Tyr Val Ile Gly Ile Ala Gly Ser Val Ala Val Gly Lys Ser Thr  
 85 90 95 100

gtt gcc agg ctt ctt aag tcc ctc ttg agc gac tac tat ccg gaa aaa 393  
 Val Ala Arg Leu Leu Lys Ser Leu Leu Ser Asp Tyr Tyr Pro Glu Lys  
 105 110 115

aag gta gac ctc ctc aca aca gat ggc ttc ctt tat ccg aat aag att 441  
 Lys Val Asp Leu Leu Thr Thr Asp Gly Phe Leu Tyr Pro Asn Lys Ile  
 120 125 130

tta aaa gag cga gat atc atg gac cgc aag ggt ttt ccc gaa agc tat 489  
 Leu Lys Glu Arg Asp Ile Met Asp Arg Lys Gly Phe Pro Glu Ser Tyr  
 135 140 145

gat atg aaa cgt ttg att aac ttt atg acc gat gtc aaa aat aat gtt 537  
 Asp Met Lys Arg Leu Ile Asn Phe Met Thr Asp Val Lys Asn Asn Val  
 150 155 160

ccc aac atc cag gtg ccc aag tat tcc cac caa gtt tac gac ata gta 585  
 Pro Asn Ile Gln Val Pro Lys Tyr Ser His Gln Val Tyr Asp Ile Val  
 165 170 175 180

gaa ggg gaa agg ttg acc att aac cag cca gac atc ttg att gtc gaa 633  
 Glu Gly Glu Arg Leu Thr Ile Asn Gln Pro Asp Ile Leu Ile Val Glu  
 185 190 195

ggg atc aat gtg ctc caa ctt cct tct aat gag aag att ttt gtt agc 681  
 Gly Ile Asn Val Leu Gln Leu Pro Ser Asn Glu Lys Ile Phe Val Ser  
 200 205 210  
 gat ttt ttc gac ttc tcc ttt tat gtg gat gcc tca gaa aat ctg att 729  
 Asp Phe Phe Asp Phe Ser Phe Tyr Val Asp Ala Ser Glu Asn Leu Ile  
 215 220 225  
 gaa aaa tgg tac atg caa cgc ttt ggc acc ttt atg gat acc gcc ttc 777  
 Glu Lys Trp Tyr Met Gln Arg Phe Gly Thr Phe Met Asp Thr Ala Phe  
 230 235 240  
 caa gac ccc aac aac tat tac tac aag ttt aat gac tgg gac cgc aag 825  
 Gln Asp Pro Asn Asn Tyr Tyr Tyr Lys Phe Asn Asp Trp Asp Arg Lys  
 245 250 255 260  
 gaa gct ttt gcc tat gcc aac caa gtt tgg gaa acg gtt aac cta gaa 873  
 Glu Ala Phe Ala Tyr Ala Asn Gln Val Trp Glu Thr Val Asn Leu Glu  
 265 270 275  
 aac ctc agg gaa tat att cta ccc acc cga ctc cgg gct aac ctc atc 921  
 Asn Leu Arg Glu Tyr Ile Leu Pro Thr Arg Leu Arg Ala Asn Leu Ile  
 280 285 290  
 ctc cat aaa acc cat aac cac tac atc gac aag att tta ctc aaa aaa 969  
 Leu His Lys Thr His Asn His Tyr Ile Asp Lys Ile Leu Leu Lys Lys  
 295 300 305  
 cac tga 975  
 His

&lt;210&gt; 48

&lt;211&gt; 309

&lt;212&gt; PRT

&lt;213&gt; Alloiococcus otitidis

&lt;400&gt; 48

Met Asn Asn Ser Arg Ile Phe Leu Val Tyr Asp Arg Lys Asp Trp Gln  
 1 5 10 15

Ser Leu Arg Glu Asn Ala Ser Leu Ser Leu Thr Glu Lys Asn Leu Asn  
 20 25 30

Asn Leu Arg Ala Val Asn Asp Val Ile Ser Met Glu Asp Val Arg Glu  
 35 40 45

Val Tyr Val Pro Ile Ile Gln Leu Leu Asp Val Tyr Ile Lys Ser Tyr  
 50 55 60

Tyr Arg His Gln Ala Ser Leu Ile Asn Tyr Leu Asn Leu Asp Gln Pro  
 65 70 75 80

Lys Lys Tyr Gln Pro Tyr Val Ile Gly Ile Ala Gly Ser Val Ala Val  
85 90 95

Gly Lys Ser Thr Val Ala Arg Leu Leu Lys Ser Leu Leu Ser Asp Tyr  
100 105 110

Tyr Pro Glu Lys Lys Val Asp Leu Leu Thr Thr Asp Gly Phe Leu Tyr  
115 120 125

Pro Asn Lys Ile Leu Lys Glu Arg Asp Ile Met Asp Arg Lys Gly Phe  
130 135 140

Pro Glu Ser Tyr Asp Met Lys Arg Leu Ile Asn Phe Met Thr Asp Val  
145 150 155 160

Lys Asn Asn Val Pro Asn Ile Gln Val Pro Lys Tyr Ser His Gln Val  
165 170 175

Tyr Asp Ile Val Glu Gly Glu Arg Leu Thr Ile Asn Gln Pro Asp Ile  
180 185 190

Leu Ile Val Glu Gly Ile Asn Val Leu Gln Leu Pro Ser Asn Glu Lys  
195 200 205

Ile Phe Val Ser Asp Phe Phe Asp Phe Ser Phe Tyr Val Asp Ala Ser  
210 215 220

Glu Asn Leu Ile Glu Lys Trp Tyr Met Gln Arg Phe Gly Thr Phe Met  
225 230 235 240

Asp Thr Ala Phe Gln Asp Pro Asn Asn Tyr Tyr Tyr Lys Phe Asn Asp  
245 250 255

Trp Asp Arg Lys Glu Ala Phe Ala Tyr Ala Asn Gln Val Trp Glu Thr  
260 265 270

Val Asn Leu Glu Asn Leu Arg Glu Tyr Ile Leu Pro Thr Arg Leu Arg  
275 280 285

Ala Asn Leu Ile Leu His Lys Thr His Asn His Tyr Ile Asp Lys Ile  
290 295 300



Leu Leu Lys Lys His  
305

<210> 49

<211> 846

<212> DNA

<213> *Alloiococcus otitidis*

<220>

<221> CDS

<222> (7)..(846)

<223>

<400> 49

agctta atg gga gac gat tta aga gaa gaa att ctt gac cga atg aag	48
Met Gly Asp Asp Leu Arg Glu Glu Ile Leu Asp Arg Met Lys	
1 5 10	
gtc caa gcc caa att aac ccc aat gag gaa att cgc cgg acc att gac	96
Val Gln Ala Gln Ile Asn Pro Asn Glu Glu Ile Arg Arg Thr Ile Asp	
15 20 25 30	
ttt atc aag gac tat ctc cag gcc cac ccc ttc ttt gaa tcc tta atc	144
Phe Ile Lys Asp Tyr Leu Gln Ala His Pro Phe Phe Glu Ser Leu Ile	
35 40 45	
ttg ggc atc tcc ggt ggc cag gat tcc acc ctc ctg ggt aag cta gcc	192
Leu Gly Ile Ser Gly Gly Gln Asp Ser Thr Leu Leu Gly Lys Leu Ala	
50 55 60	
cag atg gcc tgc ctt gaa ctg agg gaa gag gag ggg tct gac aag cca	240
Gln Met Ala Cys Leu Glu Leu Arg Glu Glu Glu Gly Ser Asp Lys Pro	
65 70 75	
att ttt att ggt atc cgc cta cct tat ggg gat caa ttt gat gaa gca	288
Ile Phe Ile Gly Ile Arg Leu Pro Tyr Gly Asp Gln Phe Asp Glu Ala	
80 85 90	
gaa gcc cag caa gcc ctc aat tgg atc cag cct gac cag gct ctg acc	336
Glu Ala Gln Gln Ala Leu Asn Trp Ile Gln Pro Asp Gln Ala Leu Thr	
95 100 105 110	
att aat atc aaa gag tcc gtt gat ggc ctg gtt gac act ttg gcc ggc	384
Ile Asn Ile Lys Glu Ser Val Asp Gly Leu Val Asp Thr Leu Ala Gly	
115 120 125	
caa ggc att gaa gtt tct gac ttt aac aag ggc aat atc aaa gct cgg	432
Gln Gly Ile Glu Val Ser Asp Phe Asn Lys Gly Asn Ile Lys Ala Arg	
130 135 140	
atc cga atg gtg gcc caa tat ggc gta gcg ggt cac ttc cac ggg gcg	480
Ile Arg Met Val Ala Gln Tyr Gly Val Ala Gly His Phe His Gly Ala	
145 150 155	
gtg tta gga tct gac cat tca gcc gaa aat gta act ggc ttt ttc acc	528

Val	Leu	Gly	Ser	Asp	His	Ser	Ala	Glu	Asn	Val	Thr	Gly	Phe	Phe	Thr		
160						165					170						
aag	cat	ggg	gac	ggc	gct	agt	gac	ctc	aac	cct	ctt	ttc	cgc	cta	aat	576	
Lys	His	Gly	Asp	Gly	Ala	Ser	Asp	Leu	Asn	Pro	Leu	Phe	Arg	Leu	Asn		
175				180					185						190		
aaa	cgt	cag	gga	cgg	gcc	ctg	ctt	gag	gaa	tta	ggg	tcc	cct	aag	aac	624	
Lys	Arg	Gln	Gly	Arg	Ala	Leu	Leu	Glu	Glu	Leu	Gly	Ser	Pro	Lys	Asn		
				195				200						205			
ttg	tac	caa	aag	acc	ccc	aca	gct	gat	ttg	gaa	gaa	gac	cag	ccc	ggc	672	
Leu	Tyr	Gln	Lys	Thr	Pro	Thr	Ala	Asp	Leu	Glu	Glu	Asp	Gln	Pro	Gly		
		210						215					220				
ttg	tca	gat	gaa	gac	aag	tta	ggg	gtt	tct	tat	gaa	gcc	att	gat	gac	720	
Leu	Ser	Asp	Glu	Asp	Lys	Leu	Gly	Val	Ser	Tyr	Glu	Ala	Ile	Asp	Asp		
		225					230					235					
tac	ttg	gag	ggc	aag	cca	gtt	agc	cag	gag	gac	cag	gca	acc	atc	gaa	768	
Tyr	Leu	Glu	Gly	Lys	Pro	Val	Ser	Gln	Glu	Asp	Gln	Ala	Thr	Ile	Glu		
	240					245					250						
aaa	tgg	tat	caa	caa	acg	gcc	cac	aag	cgc	cac	ttg	ccg	gtg	act	atc	816	
Lys	Trp	Tyr	Gln	Gln	Thr	Ala	His	Lys	Arg	His	Leu	Pro	Val	Thr	Ile		
255				260				265						270			
ttt	gat	gat	ttt	tgg	aaa	gaa	aaa	aat	tag							846	
Phe	Asp	Asp	Phe	Trp	Lys	Glu	Lys	Asn									
				275													

&lt;210&gt; 50

&lt;211&gt; 279

&lt;212&gt; PRT

&lt;213&gt; Alloiococcus otitidis

&lt;400&gt; 50

Met	Gly	Asp	Asp	Leu	Arg	Glu	Glu	Ile	Leu	Asp	Arg	Met	Lys	Val	Gln
1				5					10				15		

Ala	Gln	Ile	Asn	Pro	Asn	Glu	Glu	Ile	Arg	Arg	Thr	Ile	Asp	Phe	Ile
		20						25					30		

Lys	Asp	Tyr	Leu	Gln	Ala	His	Pro	Phe	Phe	Glu	Ser	Leu	Ile	Leu	Gly
		35						40				45			

Ile	Ser	Gly	Gly	Gln	Asp	Ser	Thr	Leu	Leu	Gly	Lys	Leu	Ala	Gln	Met
	50					55					60				

Ala	Cys	Leu	Glu	Leu	Arg	Glu	Glu	Glu	Gly	Ser	Asp	Lys	Pro	Ile	Phe
65					70					75					80

Ile Gly Ile Arg Leu Pro Tyr Gly Asp Gln Phe Asp Glu Ala Glu Ala  
85 90 95

Gln Gln Ala Leu Asn Trp Ile Gln Pro Asp Gln Ala Leu Thr Ile Asn  
100 105 110

Ile Lys Glu Ser Val Asp Gly Leu Val Asp Thr Leu Ala Gly Gln Gly  
115 120 125

Ile Glu Val Ser Asp Phe Asn Lys Gly Asn Ile Lys Ala Arg Ile Arg  
130 135 140

Met Val Ala Gln Tyr Gly Val Ala Gly His Phe His Gly Ala Val Leu  
145 150 155 160

Gly Ser Asp His Ser Ala Glu Asn Val Thr Gly Phe Phe Thr Lys His  
165 170 175

Gly Asp Gly Ala Ser Asp Leu Asn Pro Leu Phe Arg Leu Asn Lys Arg  
180 185 190

Gln Gly Arg Ala Leu Leu Glu Glu Leu Gly Ser Pro Lys Asn Leu Tyr  
195 200 205

Gln Lys Thr Pro Thr Ala Asp Leu Glu Glu Asp Gln Pro Gly Leu Ser  
210 215 220

Asp Glu Asp Lys Leu Gly Val Ser Tyr Glu Ala Ile Asp Asp Tyr Leu  
225 230 235 240

Glu Gly Lys Pro Val Ser Gln Glu Asp Gln Ala Thr Ile Glu Lys Trp  
245 250 255

Tyr Gln Gln Thr Ala His Lys Arg His Leu Pro Val Thr Ile Phe Asp  
260 265 270

Asp Phe Trp Lys Glu Lys Asn  
275

<210> 51

<211> 843

<212> DNA

<213> *Alloiococcus otitidis*

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (7)..(843)

&lt;223&gt;

&lt;400&gt; 51

aggaac atg att atg tat act gat ggg att ggc ttt att gat tca gga	48
Met Ile Met Tyr Thr Asp Gly Ile Gly Phe Ile Asp Ser Gly	
1 5 10	
gtg ggt ggc ttc acc ctg gtc aaa gaa gcc atg aag caa ttg cca aat	96
Val Gly Gly Phe Thr Leu Val Lys Glu Ala Met Lys Gln Leu Pro Asn	
15 20 25 30	
gaa caa ttt tac tat ctg gga gac acc gcc cgg tca cct tat gga cct	144
Glu Gln Phe Tyr Tyr Leu Gly Asp Thr Ala Arg Ser Pro Tyr Gly Pro	
35 40 45	
aaa gac atg gcc act gtc aag gca tat gcc ttt gaa ctt gcc aat tac	192
Lys Asp Met Ala Thr Val Lys Ala Tyr Ala Phe Glu Leu Ala Asn Tyr	
50 55 60	
ctg gtt aaa aac cac cag atc aaa atc ttg gtg atc gct tgt aat act	240
Leu Val Lys Asn His Gln Ile Lys Ile Leu Val Ile Ala Cys Asn Thr	
65 70 75	
gcg act gtc gct gcc ctc aag gac cta aaa cag gcc ttg ccc atc cca	288
Ala Thr Val Ala Ala Leu Lys Asp Leu Lys Gln Ala Leu Pro Ile Pro	
80 85 90	
gtt tta ggg gtc atc tta cct ggt tgc cga gca gct att aag gct agt	336
Val Leu Gly Val Ile Leu Pro Gly Cys Arg Ala Ala Ile Lys Ala Ser	
95 100 105 110	
gtt aac cat cag att ggg gtt att gcc acc cat ggg acc atc cag tcc	384
Val Asn His Gln Ile Gly Val Ile Ala Thr His Gly Thr Ile Gln Ser	
115 120 125	
ggg cgc tat gag ctt gaa ctt aaa cgg aaa cga ccg gat att gaa gtg	432
Gly Arg Tyr Glu Leu Glu Leu Lys Arg Lys Arg Pro Asp Ile Glu Val	
130 135 140	
aca agt ctg gct tgt ccc gaa ttt gcc ccc atg gta gag gcg gga gac	480
Thr Ser Leu Ala Cys Pro Glu Phe Ala Pro Met Val Glu Ala Gly Asp	
145 150 155	
tac cga tct gtt caa gct agc agt gtg gtg agg aca tcc tta cag gcc	528
Tyr Arg Ser Val Gln Ala Ser Ser Val Val Arg Thr Ser Leu Gln Ala	
160 165 170	
cta gaa gac caa gat ttg gat acc ctt att ttg ggt tgc acc cac tat	576
Leu Glu Asp Gln Asp Leu Asp Thr Leu Ile Leu Gly Cys Thr His Tyr	
175 180 185 190	
ccc att ata aaa gac ctc att caa gac tct att ggc cct ggt atc agc	624
Pro Ile Ile Lys Asp Leu Ile Gln Asp Ser Ile Gly Pro Gly Ile Ser	

195	200	205	
ttg gtt gat cca ggg gcg gaa gct gtg aat gac ttg agt gtc tta tta			672
Leu Val Asp Pro Gly Ala Glu Ala Val Asn Asp Leu Ser Val Leu Leu			
210	215	220	
gac tat tat gac ttg act aat gac cgg ttt aat ccc aac ctg acc cac			720
Asp Tyr Tyr Asp Leu Thr Asn Asp Arg Phe Asn Pro Asn Leu Thr His			
225	230	235	
cat ttt tac acc acg gga gat aaa gcc ggg ttt aag aaa atc gcg gat			768
His Phe Tyr Thr Thr Gly Asp Lys Ala Gly Phe Lys Lys Ile Ala Asp			
240	245	250	
gac tgg ctt gac cac cac aac tac cgg gtt gac cat tta gat tta gag			816
Asp Trp Leu Asp His His Asn Tyr Arg Val Asp His Leu Asp Leu Glu			
255	260	265	270
gag ttg caa gaa gtt aat gga aga taa			843
Glu Leu Gln Glu Val Asn Gly Arg			
275			

&lt;210&gt; 52

&lt;211&gt; 278

&lt;212&gt; PRT

&lt;213&gt; Alloioococcus otitidis

&lt;400&gt; 52

Met Ile Met Tyr Thr Asp Gly Ile Gly Phe Ile Asp Ser Gly Val Gly
1 5 10 15

Gly Phe Thr Leu Val Lys Glu Ala Met Lys Gln Leu Pro Asn Glu Gln
20 25 30

Phe Tyr Tyr Leu Gly Asp Thr Ala Arg Ser Pro Tyr Gly Pro Lys Asp
35 40 45

Met Ala Thr Val Lys Ala Tyr Ala Phe Glu Leu Ala Asn Tyr Leu Val
50 55 60

Lys Asn His Gln Ile Lys Ile Leu Val Ile Ala Cys Asn Thr Ala Thr
65 70 75 80

Val Ala Ala Leu Lys Asp Leu Lys Gln Ala Leu Pro Ile Pro Val Leu
85 90 95

Gly Val Ile Leu Pro Gly Cys Arg Ala Ala Ile Lys Ala Ser Val Asn
100 105 110

His Gln Ile Gly Val Ile Ala Thr His Gly Thr Ile Gln Ser Gly Arg  
 115 120 125

Tyr Glu Leu Glu Leu Lys Arg Lys Arg Pro Asp Ile Glu Val Thr Ser  
 130 135 140

Leu Ala Cys Pro Glu Phe Ala Pro Met Val Glu Ala Gly Asp Tyr Arg  
 145 150 155 160

Ser Val Gln Ala Ser Ser Val Val Arg Thr Ser Leu Gln Ala Leu Glu  
 165 170 175

Asp Gln Asp Leu Asp Thr Leu Ile Leu Gly Cys Thr His Tyr Pro Ile  
 180 185 190

Ile Lys Asp Leu Ile Gln Asp Ser Ile Gly Pro Gly Ile Ser Leu Val  
 195 200 205

Asp Pro Gly Ala Glu Ala Val Asn Asp Leu Ser Val Leu Leu Asp Tyr  
 210 215 220

Tyr Asp Leu Thr Asn Asp Arg Phe Asn Pro Asn Leu Thr His His Phe  
 225 230 235 240

Tyr Thr Thr Gly Asp Lys Ala Gly Phe Lys Lys Ile Ala Asp Asp Trp  
 245 250 255

Leu Asp His His Asn Tyr Arg Val Asp His Leu Asp Leu Glu Glu Leu  
 260 265 270

Gln Glu Val Asn Gly Arg  
 275

<210> 53

<211> 957

<212> DNA

<213> *Alloiococcus otitidis*

<220>

<221> CDS

<222> (7)..(957)

<223>

<400> 53

aaaaat atg acg aag gag tct tca ttt atg gtc aag acc aaa ata tgt  
 Met Thr Lys Glu Ser Ser Phe Met Val Lys Thr Lys Ile Cys

1	5	10	
tct att tta aat ata aca ccg gat tca ttt tct gat ggt ggg cgc aac Ser Ile Leu Asn Ile Thr Pro Asp Ser Phe Ser Asp Gly Gly Arg Asn 15 20 25 30	96		
tat cag gca gac caa gcc ata gct cac gga ctc gac ttg gta gac aag Tyr Gln Ala Asp Gln Ala Ile Ala His Gly Leu Asp Leu Val Asp Lys 35 40 45	144		
gga gcg gac atg ttg gat att gga ggt gag tcg acc cgg cct ggt tcc Gly Ala Asp Met Leu Asp Ile Gly Gly Glu Ser Thr Arg Pro Gly Ser 50 55 60	192		
agt cca gtc gac ctc caa gat gaa atc gac cgt att gta ccg gtg atc Ser Pro Val Asp Leu Gln Asp Glu Ile Asp Arg Ile Val Pro Val Ile 65 70 75	240		
aag gga atc aga gaa aaa agt cag gtt cct att tca gta gat acc tac Lys Gly Ile Arg Glu Lys Ser Gln Val Pro Ile Ser Val Asp Thr Tyr 80 85 90	288		
cgg gct cca gtt gcc aaa gcg gct att gat gct ggg gcg gat atc atc Arg Ala Pro Val Ala Lys Ala Ala Ile Asp Ala Gly Ala Asp Ile Ile 95 100 105 110	336		
aat gat att acc ggt cta act ggt gat gta gac atg gcc gac ttg cta Asn Asp Ile Thr Gly Leu Thr Gly Asp Val Asp Met Ala Asp Leu Leu 115 120 125	384		
gct caa gaa ggg gtt aag gcc att gtc atg ttc aac ccg gtt att gct Ala Gln Glu Gly Val Lys Ala Ile Val Met Phe Asn Pro Val Ile Ala 130 135 140	432		
cga cct gac cac cca tct tcc caa aaa ttc aga gat ttc ggg ggc cga Arg Pro Asp His Pro Ser Ser Gln Lys Phe Arg Asp Phe Gly Gly Arg 145 150 155	480		
gat ttt ttc acc gat gaa gaa aga gat aaa atg tcc caa gca ccc att Asp Phe Phe Thr Asp Glu Glu Arg Asp Lys Met Ser Gln Ala Pro Ile 160 165 170	528		
gaa gag gcc atg atg gtc tac ttt gac aaa gtc ttg aac aag gcc cat Glu Glu Ala Met Met Val Tyr Phe Asp Lys Val Leu Asn Lys Ala His 175 180 185 190	576		
caa gct ggg att gac cgg gat aag att tta ctg gac ccg gga att ggc Gln Ala Gly Ile Asp Arg Asp Lys Ile Leu Leu Asp Pro Gly Ile Gly 195 200 205	624		
ttt ggc ctg acc aag aag gaa aat tac aag ttg att cac agt gtt gcc Phe Gly Leu Thr Lys Lys Glu Asn Tyr Lys Leu Ile His Ser Val Ala 210 215 220	672		
tcg att cat gac aag ggc tac ccg gtc ttt tta gga gtt tcc cgc aaa Ser Ile His Asp Lys Gly Tyr Pro Val Phe Leu Gly Val Ser Arg Lys 225 230 235	720		

cgc ttc ttg gtg ggg gaa gtc tcc aag cta ggc atc gaa gcc gac cca 768  
 Arg Phe Leu Val Gly Glu Val Ser Lys Leu Gly Ile Glu Ala Asp Pro  
 240 245 250

gag acc caa gca gga ttt tta aac cga gac ctg gct tca gct att att 816  
 Glu Thr Gln Ala Gly Phe Leu Asn Arg Asp Leu Ala Ser Ala Ile Ile  
 255 260 265 270

aca gct tac gct agc cat ata ggg gta gac tat gtc cgg gtt cat tcc 864  
 Thr Ala Tyr Ala Ser His Ile Gly Val Asp Tyr Val Arg Val His Ser  
 275 280 285

tta gat gaa cac aaa ata gca acc acc att acc cat aat att tta aac 912  
 Leu Asp Glu His Lys Ile Ala Thr Thr Ile Thr His Asn Ile Leu Asn  
 290 295 300

agc gat agc tta gat gat cag agc ttt gac caa tat aaa aat taa 957  
 Ser Asp Ser Leu Asp Asp Gln Ser Phe Asp Gln Tyr Lys Asn  
 305 310 315

<210> 54

<211> 316

<212> PRT

<213> Alloiococcus otitidis

<400> 54

Met Thr Lys Glu Ser Ser Phe Met Val Lys Thr Lys Ile Cys Ser Ile  
 1 5 10 15

Leu Asn Ile Thr Pro Asp Ser Phe Ser Asp Gly Gly Arg Asn Tyr Gln  
 20 25 30

Ala Asp Gln Ala Ile Ala His Gly Leu Asp Leu Val Asp Lys Gly Ala  
 35 40 45

Asp Met Leu Asp Ile Gly Gly Glu Ser Thr Arg Pro Gly Ser Ser Pro  
 50 55 60

Val Asp Leu Gln Asp Glu Ile Asp Arg Ile Val Pro Val Ile Lys Gly  
 65 70 75 80

Ile Arg Glu Lys Ser Gln Val Pro Ile Ser Val Asp Thr Tyr Arg Ala  
 85 90 95

Pro Val Ala Lys Ala Ala Ile Asp Ala Gly Ala Asp Ile Ile Asn Asp  
 100 105 110

Ile Thr Gly Leu Thr Gly Asp Val Asp Met Ala Asp Leu Leu Ala Gln



115

120

125

Glu Gly Val Lys Ala Ile Val Met Phe Asn Pro Val Ile Ala Arg Pro  
 130 135 140

Asp His Pro Ser Ser Gln Lys Phe Arg Asp Phe Gly Gly Arg Asp Phe  
 145 150 155 160

Phe Thr Asp Glu Glu Arg Asp Lys Met Ser Gln Ala Pro Ile Glu Glu  
 165 170 175

Ala Met Met Val Tyr Phe Asp Lys Val Leu Asn Lys Ala His Gln Ala  
 180 185 190

Gly Ile Asp Arg Asp Lys Ile Leu Leu Asp Pro Gly Ile Gly Phe Gly  
 195 200 205

Leu Thr Lys Lys Glu Asn Tyr Lys Leu Ile His Ser Val Ala Ser Ile  
 210 215 220

His Asp Lys Gly Tyr Pro Val Phe Leu Gly Val Ser Arg Lys Arg Phe  
 225 230 235 240

Leu Val Gly Glu Val Ser Lys Leu Gly Ile Glu Ala Asp Pro Glu Thr  
 245 250 255

Gln Ala Gly Phe Leu Asn Arg Asp Leu Ala Ser Ala Ile Ile Thr Ala  
 260 265 270

Tyr Ala Ser His Ile Gly Val Asp Tyr Val Arg Val His Ser Leu Asp  
 275 280 285

Glu His Lys Ile Ala Thr Thr Ile Thr His Asn Ile Leu Asn Ser Asp  
 290 295 300

Ser Leu Asp Asp Gln Ser Phe Asp Gln Tyr Lys Asn  
 305 310 315

&lt;210&gt; 55

&lt;211&gt; 561

&lt;212&gt; DNA

<213> *Alloiococcus otitidis*

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (28)..(561)

&lt;223&gt;

&lt;400&gt; 55

acaagaacta gacgaaaaat tagaacg ttg gga ata ttt aag cca ata tgt ata 54  
 Met Gly Ile Phe Lys Pro Ile Cys Ile  
 1 5

gga gag ata act atg ata gcc tac gtt tgg gcc caa gat gag caa gga 102  
 Gly Glu Ile Thr Met Ile Ala Tyr Val Trp Ala Gln Asp Glu Gln Gly  
 10 15 20 25

atc att ggt aaa gac aag gtt ttg cct tgg gaa ttg tcc aat gac tta 150  
 Ile Ile Gly Lys Asp Lys Val Leu Pro Trp Glu Leu Ser Asn Asp Leu  
 30 35 40

aag cat ttt aaa aaa gtt aca gaa ggt cac acc atc ctg atg ggc cgg 198  
 Lys His Phe Lys Lys Val Thr Glu Gly His Thr Ile Leu Met Gly Arg  
 45 50 55

aag acc ttt gaa gga atg gat aaa aag ccc ctc cct aac cga aaa acc 246  
 Lys Thr Phe Glu Gly Met Asp Lys Lys Pro Leu Pro Asn Arg Lys Thr  
 60 65 70

ttg gta ttg acc cgc caa gat gac tac caa gct ggg gac gac cag gtt 294  
 Leu Val Leu Thr Arg Gln Asp Asp Tyr Gln Ala Gly Asp Asp Gln Val  
 75 80 85

gaa gtc gtc cac tcc aaa gac cag gcc ttg act tat gcg tca ggt cat 342  
 Glu Val Val His Ser Lys Asp Gln Ala Leu Thr Tyr Ala Ser Gly His  
 90 95 100 105

ggg gtg gac ctc tat gtg att ggt ggg gcc ggc att ttc gac ttg ttt 390  
 Gly Val Asp Leu Tyr Val Ile Gly Gly Ala Gly Ile Phe Asp Leu Phe  
 110 115 120

ctg gac caa gtt gat gtt ctc cac caa aca gtt atc cac gag agc ttt 438  
 Leu Asp Gln Val Asp Val Leu His Gln Thr Val Ile His Glu Ser Phe  
 125 130 135

gat ggt gac acc acc atg cca gac att gac tgg gac agc ttt aat cag 486  
 Asp Gly Asp Thr Thr Met Pro Asp Ile Asp Trp Asp Ser Phe Asn Gln  
 140 145 150

gtg tct aaa gct tat tat gac cag gct gac ggt cac aac cac tcc cac 534  
 Val Ser Lys Ala Tyr Tyr Asp Gln Ala Asp Gly His Asn His Ser His  
 155 160 165

acc att tat gaa tac aga aga aaa taa 561  
 Thr Ile Tyr Glu Tyr Arg Arg Lys  
 170 175

&lt;210&gt; 56

&lt;211&gt; 177

&lt;212&gt; PRT

<213> *Alloiococcus otitidis*

<400> 56

Met Gly Ile Phe Lys Pro Ile Cys Ile Gly Glu Ile Thr Met Ile Ala  
1 5 10 15

Tyr Val Trp Ala Gln Asp Glu Gln Gly Ile Ile Gly Lys Asp Lys Val  
20 25 30

Leu Pro Trp Glu Leu Ser Asn Asp Leu Lys His Phe Lys Lys Val Thr  
35 40 45

Glu Gly His Thr Ile Leu Met Gly Arg Lys Thr Phe Glu Gly Met Asp  
50 55 60

Lys Lys Pro Leu Pro Asn Arg Lys Thr Leu Val Leu Thr Arg Gln Asp  
65 70 75 80

Asp Tyr Gln Ala Gly Asp Asp Gln Val Glu Val Val His Ser Lys Asp  
85 90 95

Gln Ala Leu Thr Tyr Ala Ser Gly His Gly Val Asp Leu Tyr Val Ile  
100 105 110

Gly Gly Ala Gly Ile Phe Asp Leu Phe Leu Asp Gln Val Asp Val Leu  
115 120 125

His Gln Thr Val Ile His Glu Ser Phe Asp Gly Asp Thr Thr Met Pro  
130 135 140

Asp Ile Asp Trp Asp Ser Phe Asn Gln Val Ser Lys Ala Tyr Tyr Asp  
145 150 155 160

Gln Ala Asp Gly His Asn His Ser His Thr Ile Tyr Glu Tyr Arg Arg  
165 170 175

Lys

<210> 57

<211> 1968

<212> DNA

<213> *Alloiococcus otitidis*

<220>

<221> CDS

&lt;222&gt; (7) .. (1968)

&lt;223&gt;

&lt;400&gt; 57

agagat ttg acg aag gaa tct tac aat gat tca tct ata acc ata ctc	48
Met Thr Lys Glu Ser Tyr Asn Asp Ser Ser Ile Thr Ile Leu	
1 5 10	
aag ggc tta gac gcc gtt aag aaa aga cca ggc atg tat atc ggg tca	96
Lys Gly Leu Asp Ala Val Lys Lys Arg Pro Gly Met Tyr Ile Gly Ser	
15 20 25 30	
acc gat gcc agg ggt ttg cac cac ctg gtt tat gaa att acc gat aat	144
Thr Asp Ala Arg Gly Leu His His Leu Val Tyr Glu Ile Thr Asp Asn	
35 40 45	
gct att gat gag gtt ttg gct ggc tac gct gat gaa att gaa gtc aag	192
Ala Ile Asp Glu Val Leu Ala Gly Tyr Ala Asp Glu Ile Glu Val Lys	
50 55 60	
atc cac acg gac ggc tcg gtt tcg gtc aaa gac aat gga cgg ggc atg	240
Ile His Thr Asp Gly Ser Val Ser Val Lys Asp Asn Gly Arg Gly Met	
65 70 75	
cca acc ggg atg cat gag tca ggc cta ccc acc atc cag gtt atc ttt	288
Pro Thr Gly Met His Glu Ser Gly Leu Pro Thr Ile Gln Val Ile Phe	
80 85 90	
acc gtc ctc cat gcc ggg gga aaa ttt ggc caa gag ggg gcc tac aag	336
Thr Val Leu His Ala Gly Gly Lys Phe Gly Gln Glu Gly Ala Tyr Lys	
95 100 105 110	
tca gcc ggt gga ctc cat ggg gtt ggg gcc tcg gtc gtc aac gcc ttg	384
Ser Ala Gly Gly Leu His Gly Val Gly Ala Ser Val Val Asn Ala Leu	
115 120 125	
tct gat tgg ctc acg gtg ata gtg acc aag gac ggc tat gaa tac cgg	432
Ser Asp Trp Leu Thr Val Ile Val Thr Lys Asp Gly Tyr Glu Tyr Arg	
130 135 140	
caa gac ttt agc caa gga ggc cag gct aaa gga ggc atc cag aag aga	480
Gln Asp Phe Ser Gln Gly Gly Gln Ala Lys Gly Gly Ile Gln Lys Arg	
145 150 155	
aaa att aac cag caa aaa tcc agc acc ctg gtc cac ttc aaa ccc tca	528
Lys Ile Asn Gln Gln Lys Ser Ser Thr Leu Val His Phe Lys Pro Ser	
160 165 170	
ggc caa gtc ttt tcg acc acc gaa ttt aac ttt aac acc atc tgt gag	576
Gly Gln Val Phe Ser Thr Thr Glu Phe Asn Phe Asn Thr Ile Cys Glu	
175 180 185 190	
cgg atg cgg gag tcg gcc ttc ctt gtc aaa ggg acc aag att acc gta	624
Arg Met Arg Glu Ser Ala Phe Leu Val Lys Gly Thr Lys Ile Thr Val	
195 200 205	
gag gac ctg cgc cag gaa gaa agc cag gtc ttc caa ttt aat gaa gga	672

Glu Asp Leu Arg Gln Glu Glu Ser Gln Val Phe Gln Phe Asn Glu Gly	
210 215 220	
att aag gcc ttt gtc gac tac tta aat gag ggc aag gat acc ttg agt	720
Ile Lys Ala Phe Val Asp Tyr Leu Asn Glu Gly Lys Asp Thr Leu Ser	
225 230 235	
cca gta acc tat ttt gaa ggt tct gaa gat gaa att gaa gtt gaa ttt	768
Pro Val Thr Tyr Phe Glu Gly Ser Glu Asp Glu Ile Glu Val Glu Phe	
240 245 250	
gcc ttc caa tac aat gac ggc tat tct gag acg gtt ctg agt ttt gtc	816
Ala Phe Gln Tyr Asn Asp Gly Tyr Ser Glu Thr Val Leu Ser Phe Val	
255 260 265 270	
aac aat gtc cgt acc cgg gat ggg ggc agc cac gaa act gga gct aag	864
Asn Asn Val Arg Thr Arg Asp Gly Gly Ser His Glu Thr Gly Ala Lys	
275 280 285	
tca gct att acc aag gct ttc aac gac tat gct agg aaa agt ggc tta	912
Ser Ala Ile Thr Lys Ala Phe Asn Asp Tyr Ala Arg Lys Ser Gly Leu	
290 295 300	
ctc aaa gag aaa gac agt aac ttg gaa gga tct gac gtc cgg gaa ggg	960
Leu Lys Glu Lys Asp Ser Asn Leu Glu Gly Ser Asp Val Arg Glu Gly	
305 310 315	
att gcg gtt gtt tta tcc gtc cgt atc cca gaa gag att ctc caa ttt	1008
Ile Ala Val Val Leu Ser Val Arg Ile Pro Glu Glu Ile Leu Gln Phe	
320 325 330	
gaa ggc cag acc aag agc aag tta gga act cct caa gcc cgg acc gcc	1056
Glu Gly Gln Thr Lys Ser Lys Leu Gly Thr Pro Gln Ala Arg Thr Ala	
335 340 345 350	
act gac cag gtt atc tca gaa tcc tta act tac ttc ctg gcc gaa aat	1104
Thr Asp Gln Val Ile Ser Glu Ser Leu Thr Tyr Phe Leu Ala Glu Asn	
355 360 365	
ggg gac ttg tct aag caa ctt att cgc aag gcc atc cga gcc cgg tct	1152
Gly Asp Leu Ser Lys Gln Leu Ile Arg Lys Ala Ile Arg Ala Arg Ser	
370 375 380	
gcc agg gaa gca gct cgc aag gcc aag gac cag tcc cgg aac tct gct	1200
Ala Arg Glu Ala Ala Arg Lys Ala Lys Asp Gln Ser Arg Asn Ser Ala	
385 390 395	
tcc aag aaa aaa gtt gaa act ctc ctg tct ggt aag ttg acc cca gct	1248
Ser Lys Lys Lys Val Glu Thr Leu Leu Ser Gly Lys Leu Thr Pro Ala	
400 405 410	
caa agc aag aac gcc cag aaa aat gaa ctt tac tta gtg gag ggg gat	1296
Gln Ser Lys Asn Ala Gln Lys Asn Glu Leu Tyr Leu Val Glu Gly Asp	
415 420 425 430	
tcg gct ggt ggg tca gcc aag caa ggt agg gac cgg aaa ttc caa gca	1344
Ser Ala Gly Gly Ser Ala Lys Gln Gly Arg Asp Arg Lys Phe Gln Ala	

435	440	445	
att ttg ccc ctg cgt gga aag gtt atc aac aca gaa aaa tct tct ttg Ile Leu Pro Leu Arg Gly Lys Val Ile Asn Thr Glu Lys Ser Ser Leu 450 455 460			1392
gat gat att tta aaa aat gaa gaa att tct acc atg att tat acc atc Asp Asp Ile Leu Lys Asn Glu Glu Ile Ser Thr Met Ile Tyr Thr Ile 465 470 475			1440
ggt gca ggt gct ggg cct gag ttt gat att gaa gct gtt aat tac gat Gly Ala Gly Ala Gly Pro Phe Asp Ile Glu Ala Val Asn Tyr Asp 480 485 490			1488
aag ata gtc att atg act gat gcc gac aca gac ggc gcc cac atc cag Lys Ile Val Ile Met Thr Asp Ala Asp Thr Asp Gly Ala His Ile Gln 495 500 505 510			1536
gtc ctt ctc ctc acc ttc ttt tac cgg tac atg aaa ccc ctg att gaa Val Leu Leu Leu Thr Phe Phe Tyr Arg Tyr Met Lys Pro Leu Ile Glu 515 520 525			1584
gca ggg aag gtc tat att gcc cta ccg ccc ttg tat aag ttg acc aaa Ala Gly Lys Val Tyr Ile Ala Leu Pro Pro Leu Tyr Lys Leu Thr Lys 530 535 540			1632
aag caa gga aag caa gaa aaa aca gcc tat gct tgg act gat gag gag Lys Gln Gly Lys Gln Glu Lys Thr Ala Tyr Ala Trp Thr Asp Glu Glu 545 550 555			1680
ttg gaa gac ctg gtt aaa gat ttt ggc aaa cac tac act ctc cag cgc Leu Glu Asp Leu Val Lys Asp Phe Gly Lys His Tyr Thr Leu Gln Arg 560 565 570			1728
tac aag ggt tta ggc gag atg aat gct gac cag ttg tgg gag acc acc Tyr Lys Gly Leu Gly Glu Met Asn Ala Asp Gln Leu Trp Glu Thr Thr 575 580 585 590			1776
atg gac cca gag acc aga acc ttg atc cgg gtc acc att gaa gac agt Met Asp Pro Glu Thr Arg Thr Leu Ile Arg Val Thr Ile Glu Asp Ser 595 600 605			1824
gaa aag gct gaa aga cgg gtt tcc acc ttg atg ggg acc aag gtg gat Glu Lys Ala Glu Arg Arg Val Ser Thr Leu Met Gly Thr Lys Val Asp 610 615 620			1872
cct aga cgg aag tgg att gaa gac cat att gaa ttc agt ctg gca gaa Pro Arg Arg Lys Trp Ile Glu Asp His Ile Glu Phe Ser Leu Ala Glu 625 630 635			1920
gat ggc agt att tta gag aac aag gtc cta gaa gga gag gcc aag taa Asp Gly Ser Ile Leu Glu Asn Lys Val Leu Glu Gly Glu Ala Lys 640 645 650			1968

&lt;210&gt; 58

&lt;211&gt; 653

&lt;212&gt; PRT

&lt;213&gt; Alloiococcus otitidis

&lt;400&gt; 58

Met Thr Lys Glu Ser Tyr Asn Asp Ser Ser Ile Thr Ile Leu Lys Gly  
1 5 10 15

Leu Asp Ala Val Lys Lys Arg Pro Gly Met Tyr Ile Gly Ser Thr Asp  
20 25 30

Ala Arg Gly Leu His His Leu Val Tyr Glu Ile Thr Asp Asn Ala Ile  
35 40 45

Asp Glu Val Leu Ala Gly Tyr Ala Asp Glu Ile Glu Val Lys Ile His  
50 55 60

Thr Asp Gly Ser Val Ser Val Lys Asp Asn Gly Arg Gly Met Pro Thr  
65 70 75 80

Gly Met His Glu Ser Gly Leu Pro Thr Ile Gln Val Ile Phe Thr Val  
85 90 95

Leu His Ala Gly Gly Lys Phe Gly Gln Glu Gly Ala Tyr Lys Ser Ala  
100 105 110

Gly Gly Leu His Gly Val Gly Ala Ser Val Val Asn Ala Leu Ser Asp  
115 120 125

Trp Leu Thr Val Ile Val Thr Lys Asp Gly Tyr Glu Tyr Arg Gln Asp  
130 135 140

Phe Ser Gln Gly Gly Gln Ala Lys Gly Gly Ile Gln Lys Arg Lys Ile  
145 150 155 160

Asn Gln Gln Lys Ser Ser Thr Leu Val His Phe Lys Pro Ser Gly Gln  
165 170 175

Val Phe Ser Thr Thr Glu Phe Asn Phe Asn Thr Ile Cys Glu Arg Met  
180 185 190

Arg Glu Ser Ala Phe Leu Val Lys Gly Thr Lys Ile Thr Val Glu Asp  
195 200 205

Leu Arg Gln Glu Glu Ser Gln Val Phe Gln Phe Asn Glu Gly Ile Lys

210

215

220

Ala Phe Val Asp Tyr Leu Asn Glu Gly Lys Asp Thr Leu Ser Pro Val  
225 230 235 240

Thr Tyr Phe Glu Gly Ser Glu Asp Glu Ile Glu Val Glu Phe Ala Phe  
245 250 255

Gln Tyr Asn Asp Gly Tyr Ser Glu Thr Val Leu Ser Phe Val Asn Asn  
260 265 270

Val Arg Thr Arg Asp Gly Gly Ser His Glu Thr Gly Ala Lys Ser Ala  
275 280 285

Ile Thr Lys Ala Phe Asn Asp Tyr Ala Arg Lys Ser Gly Leu Leu Lys  
290 295 300

Glu Lys Asp Ser Asn Leu Glu Gly Ser Asp Val Arg Glu Gly Ile Ala  
305 310 315 320

Val Val Leu Ser Val Arg Ile Pro Glu Glu Ile Leu Gln Phe Glu Gly  
325 330 335

Gln Thr Lys Ser Lys Leu Gly Thr Pro Gln Ala Arg Thr Ala Thr Asp  
340 345 350

Gln Val Ile Ser Glu Ser Leu Thr Tyr Phe Leu Ala Glu Asn Gly Asp  
355 360 365

Leu Ser Lys Gln Leu Ile Arg Lys Ala Ile Arg Ala Arg Ser Ala Arg  
370 375 380

Glu Ala Ala Arg Lys Ala Lys Asp Gln Ser Arg Asn Ser Ala Ser Lys  
385 390 395 400

Lys Lys Val Glu Thr Leu Leu Ser Gly Lys Leu Thr Pro Ala Gln Ser  
405 410 415

Lys Asn Ala Gln Lys Asn Glu Leu Tyr Leu Val Glu Gly Asp Ser Ala  
420 425 430

Gly Gly Ser Ala Lys Gln Gly Arg Asp Arg Lys Phe Gln Ala Ile Leu  
435 440 445



Pro Leu Arg Gly Lys Val Ile Asn Thr Glu Lys Ser Ser Leu Asp Asp  
450 455 460

Ile Leu Lys Asn Glu Glu Ile Ser Thr Met Ile Tyr Thr Ile Gly Ala  
465 470 475 480

Gly Ala Gly Pro Glu Phe Asp Ile Glu Ala Val Asn Tyr Asp Lys Ile  
485 490 495

Val Ile Met Thr Asp Ala Asp Thr Asp Gly Ala His Ile Gln Val Leu  
500 505 510

Leu Leu Thr Phe Phe Tyr Arg Tyr Met Lys Pro Leu Ile Glu Ala Gly  
515 520 525

Lys Val Tyr Ile Ala Leu Pro Pro Leu Tyr Lys Leu Thr Lys Lys Gln  
530 535 540

Gly Lys Gln Glu Lys Thr Ala Tyr Ala Trp Thr Asp Glu Glu Leu Glu  
545 550 555 560

Asp Leu Val Lys Asp Phe Gly Lys His Tyr Thr Leu Gln Arg Tyr Lys  
565 570 575

Gly Leu Gly Glu Met Asn Ala Asp Gln Leu Trp Glu Thr Thr Met Asp  
580 585 590

Pro Glu Thr Arg Thr Leu Ile Arg Val Thr Ile Glu Asp Ser Glu Lys  
595 600 605

Ala Glu Arg Arg Val Ser Thr Leu Met Gly Thr Lys Val Asp Pro Arg  
610 615 620

Arg Lys Trp Ile Glu Asp His Ile Glu Phe Ser Leu Ala Glu Asp Gly  
625 630 635 640

Ser Ile Leu Glu Asn Lys Val Leu Glu Gly Glu Ala Lys  
645 650

<210> 59

<211> 2463

<212> DNA

<213> *Alloioicoccus otitidis*

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (4)..(2463)

&lt;223&gt;

&lt;400&gt; 59

att atg gca gga gac caa gag acc agt aaa ata caa gaa tta acc tta	48
Met Ala Gly Asp Gln Glu Thr Ser Lys Ile Gln Glu Leu Thr Leu	
1 5 10 15	
gaa gat gtc atg ggg gac cgg ttc ggc cgg tat tcc aag tac att ata	96
Glu Asp Val Met Gly Asp Arg Phe Gly Arg Tyr Ser Lys Tyr Ile Ile	
20 25 30	
cag gaa agg gcc cta ccg gac ttg cgg gac ggt tta aaa ccg gtc caa	144
Gln Glu Arg Ala Leu Pro Asp Leu Arg Asp Gly Leu Lys Pro Val Gln	
35 40 45	
aga cgg atc ctc tat gcc atg cac cag gac aaa aac acc tat gac aag	192
Arg Arg Ile Leu Tyr Ala Met His Gln Asp Lys Asn Thr Tyr Asp Lys	
50 55 60	
gct tac cgg aag tcg gcc aag acg gtg gga aat gtc ata ggg aac tac	240
Ala Tyr Arg Lys Ser Ala Lys Thr Val Gly Asn Val Ile Gly Asn Tyr	
65 70 75	
cac ccc cat ggc gac aca tcc gtt tac gat gcc atg gtt agg ctc agt	288
His Pro His Gly Asp Thr Ser Val Tyr Asp Ala Met Val Arg Leu Ser	
80 85 90 95	
cag cct tgg aag atg cgc cat cct ttg gtt gat atg cac ggg aac aag	336
Gln Pro Trp Lys Met Arg His Pro Leu Val Asp Met His Gly Asn Lys	
100 105 110	
ggg agc atg gac ggg gac cca cca gct gcc atg cgg tac acc gaa gcc	384
Gly Ser Met Asp Gly Asp Pro Pro Ala Ala Met Arg Tyr Thr Glu Ala	
115 120 125	
cgt ctg tct aaa att gct tcc gac ctc ctg gct gat att gat aag gag	432
Arg Leu Ser Lys Ile Ala Ser Asp Leu Leu Ala Asp Ile Asp Lys Glu	
130 135 140	
acg gtg gac cat gtc tta aac ttt gat gac acg acc gag gag ccc acc	480
Thr Val Asp His Val Leu Asn Phe Asp Asp Thr Thr Glu Glu Pro Thr	
145 150 155	
gtc tta ccc gcc cgt ttt ccc aac ctc ttg gtc aat ggg gct agc ggg	528
Val Leu Pro Ala Arg Phe Pro Asn Leu Leu Val Asn Gly Ala Ser Gly	
160 165 170 175	
att tca gcc ggt tat gct act gac ata ccg ccc cat aat ttg agc gag	576
Ile Ser Ala Gly Tyr Ala Thr Asp Ile Pro Pro His Asn Leu Ser Glu	
180 185 190	
gtg att gat gcc acc atc cac tta atc aac cac ccc aat gca agg ctg	624

Val Ile Asp Ala Thr Ile His Leu Ile Asn His Pro Asn Ala Arg Leu	
195 200 205	
gag act ttg atg gac tat att caa gga cca gac ttt ccg act ggg ggg	672
Glu Thr Leu Met Asp Tyr Ile Gln Gly Pro Asp Phe Pro Thr Gly Gly	
210 215 220	
att atc caa ggt aaa agt ggc ctg aag aaa gcc tac caa acg ggc aag	720
Ile Ile Gln Gly Lys Ser Gly Leu Lys Lys Ala Tyr Gln Thr Gly Lys	
225 230 235	
gga aaa att atc atc cgg gcc aaa gca gat att gag gcc atc cgg ggt	768
Gly Lys Ile Ile Ile Arg Ala Lys Ala Asp Ile Glu Ala Ile Arg Gly	
240 245 250 255	
ggc aaa tcc caa att gtc atc agt caa att cct tat gag gtc aac aag	816
Gly Lys Ser Gln Ile Val Ile Ser Gln Ile Pro Tyr Glu Val Asn Lys	
260 265 270	
gca agg ttg gtc caa aaa att gac gac atc cgg att aac aaa aaa atc	864
Ala Arg Leu Val Gln Lys Ile Asp Asp Ile Arg Ile Asn Lys Lys Ile	
275 280 285	
gac ggc att gcc gat gtc cgg gat gaa agt gac cgg tct ggc ttg cgg	912
Asp Gly Ile Ala Asp Val Arg Asp Glu Ser Asp Arg Ser Gly Leu Arg	
290 295 300	
att gtg gtc gaa acc aaa aaa gat ggt gat ggg gaa ggg atc tta acc	960
Ile Val Val Glu Thr Lys Lys Asp Gly Asp Gly Glu Gly Ile Leu Thr	
305 310 315	
tac ctg ctg aaa aac acc gac ctc cag gta act tat aac tta aat atg	1008
Tyr Leu Leu Lys Asn Thr Asp Leu Gln Val Thr Tyr Asn Leu Asn Met	
320 325 330 335	
gta gcc att gat aaa aaa cga ccc cag caa gtc tcc ctc aag caa atc	1056
Val Ala Ile Asp Lys Lys Arg Pro Gln Gln Val Ser Leu Lys Gln Ile	
340 345 350	
tta tct tct tac ttg gac cac aag cgg aca gtg gtt caa aac cgg acc	1104
Leu Ser Ser Tyr Leu Asp His Lys Arg Thr Val Val Gln Asn Arg Thr	
355 360 365	
cgt tac ctc tta gcc aag gcc aag gac cgc cag cac att gtc caa ggc	1152
Arg Tyr Leu Leu Ala Lys Ala Lys Asp Arg Gln His Ile Val Gln Gly	
370 375 380	
ctt atc aag gcc att tca atc ctg gat gac ttg atc caa acc atc cgg	1200
Leu Ile Lys Ala Ile Ser Ile Leu Asp Asp Leu Ile Gln Thr Ile Arg	
385 390 395	
gcc agt gaa aac aag gcc aat gcc aag gaa aat att atc cag gct tat	1248
Ala Ser Glu Asn Lys Ala Asn Ala Lys Glu Asn Ile Ile Gln Ala Tyr	
400 405 410 415	
ggg ttt agc caa gac caa gcc gaa gcc att gtc tcc ctc cag ctt tac	1296
Gly Phe Ser Gln Asp Gln Ala Glu Ala Ile Val Ser Leu Gln Leu Tyr	

420	425	430	
cgc ttg acc aat aca gat ata aag gac tta caa gca gaa gcc aaa gac Arg Leu Thr Asn Thr Asp Ile Lys Asp Leu Gln Ala Glu Ala Lys Asp 435 440 445			1344
tta gcc caa gcc atc ctg acc tac cag gac ctc tta acc aac aag gcc Leu Ala Gln Ala Ile Leu Thr Tyr Gln Asp Leu Leu Thr Asn Lys Ala 450 455 460			1392
agc ctg gat gct ttg atg aaa gaa gaa ttg aaa gaa gtc aaa caa gca Ser Leu Asp Ala Leu Met Lys Glu Glu Leu Lys Glu Val Lys Gln Ala 465 470 475			1440
tat ggg gag gac cgg cta acc cag gtc caa gac aag atc gaa aaa cta Tyr Gly Glu Asp Arg Leu Thr Gln Val Gln Asp Lys Ile Glu Lys Leu 480 485 490 495			1488
gaa ata gaa acc caa gtc ctg gtc agt gaa gaa gac gtc atg gtt acc Glu Ile Glu Thr Gln Val Leu Val Ser Glu Glu Asp Val Met Val Thr 500 505 510			1536
gtc acc cag gga ggt tac ttg aag cgg acc tcc atc cgg tct tac aag Val Thr Gln Gly Gly Tyr Leu Lys Arg Thr Ser Ile Arg Ser Tyr Lys 515 520 525			1584
gct tcc caa gtg gag gaa ttg ggc cgg cga gaa gac gac ttg gtc atc Ala Ser Gln Val Glu Glu Leu Gly Arg Arg Glu Asp Asp Leu Val Ile 530 535 540			1632
ttt atg caa gag ttg tca acc cta gac caa ctc ctt att ttc acc tcg Phe Met Gln Glu Leu Ser Thr Leu Asp Gln Leu Leu Ile Phe Thr Ser 545 550 555			1680
aaa ggc aat gtg gtc aac cga cca gtc cat gaa tta ccg gac atc aag Lys Gly Asn Val Val Asn Arg Pro Val His Glu Leu Pro Asp Ile Lys 560 565 570 575			1728
tggaag gat att gga gag cac ttg tca agg acc atc ccc ctt gga gag Trp Lys Asp Ile Gly Glu His Leu Ser Arg Thr Ile Pro Leu Gly Glu 580 585 590			1776
gac gag gaa ttg att aag gtg tac cct tat cgg gaa tta gat gcc ggc Asp Glu Glu Leu Ile Lys Val Tyr Pro Tyr Arg Glu Leu Asp Ala Gly 595 600 605			1824
aag cgc tat gtc ttt atc act cga gat ggc tat atc aaa caa agt cca Lys Arg Tyr Val Phe Ile Thr Arg Asp Gly Tyr Ile Lys Gln Ser Pro 610 615 620			1872
gag acg gaa ttt gag ccc aaa cga act tac aag tct cgg gct tca act Glu Thr Glu Phe Glu Pro Lys Arg Thr Tyr Lys Ser Arg Ala Ser Thr 625 630 635			1920
gcc att aaa tta aaa tca gac caa gat aga ctc cag gca gtc tac tat Ala Ile Lys Leu Lys Ser Asp Gln Asp Arg Leu Gln Ala Val Tyr Tyr 640 645 650 655			1968

att cct gac caa gaa gat tac gat gta ttc cta gcc agc tac aag ggc 2016  
 Ile Pro Asp Gln Glu Asp Tyr Asp Val Phe Leu Ala Ser Tyr Lys Gly  
 660 665 670

tac ggg ctc aag tat gga cta gaa gaa gtg tca gaa gta ggg gcc cag 2064  
 Tyr Gly Leu Lys Tyr Gly Leu Glu Glu Val Ser Glu Val Gly Ala Gln  
 675 680 685

gct gca ggc gtc aag tcc atg aac ctg aaa gag ggg gac cat gtc caa 2112  
 Ala Ala Gly Val Lys Ser Met Asn Leu Lys Glu Gly Asp His Val Gln  
 690 695 700

gat ggt ttg gtc ttt aag cgt aag cag ttc caa gaa gcc ttg ttc att 2160  
 Asp Gly Leu Val Phe Lys Arg Lys Gln Phe Gln Glu Ala Leu Phe Ile  
 705 710 715

acc cag cga gcc agt gtt aag aaa atg gcc ctc cat gac ttt gac cgg 2208  
 Thr Gln Arg Ala Ser Val Lys Lys Met Ala Leu His Asp Phe Asp Arg  
 720 725 730 735

act tca cgg gcc aag cgg ggt tta caa atc ctc aga gaa ctg aag cga 2256  
 Thr Ser Arg Ala Lys Arg Gly Leu Gln Ile Leu Arg Glu Leu Lys Arg  
 740 745 750

aac ccc cac cga atc cag ttt atg atc gga att tca caa aat aaa ttc 2304  
 Asn Pro His Arg Ile Gln Phe Met Ile Gly Ile Ser Gln Asn Lys Phe  
 755 760 765

ctg gtc aat ctc cta act gat aca aaa aaa cta gta cag ata aac cca 2352  
 Leu Val Asn Leu Leu Thr Asp Thr Lys Lys Leu Val Gln Ile Asn Pro  
 770 775 780

gat gac tat aca gtt tca aac cgc cat aac aat ggg tct ttt gtc ctg 2400  
 Asp Asp Tyr Thr Val Ser Asn Arg His Asn Asn Gly Ser Phe Val Leu  
 785 790 795

gac aca agc cga gat ggc aag cct gtt tct tac tat tta agt gat aac 2448  
 Asp Thr Ser Arg Asp Gly Lys Pro Val Ser Tyr Tyr Leu Ser Asp Asn  
 800 805 810 815

gat tct cac ttg taa 2463  
 Asp Ser His Leu

&lt;210&gt; 60

&lt;211&gt; 819

&lt;212&gt; PRT

<213> *Alloiococcus otitidis*

&lt;400&gt; 60

Met Ala Gly Asp Gln Glu Thr Ser Lys Ile Gln Glu Leu Thr Leu Glu  
 1 5 10 15

Asp Val Met Gly Asp Arg Phe Gly Arg Tyr Ser Lys Tyr Ile Ile Gln  
 20 25 30

Glu Arg Ala Leu Pro Asp Leu Arg Asp Gly Leu Lys Pro Val Gln Arg  
35 40 45

Arg Ile Leu Tyr Ala Met His Gln Asp Lys Asn Thr Tyr Asp Lys Ala  
50 55 60

Tyr Arg Lys Ser Ala Lys Thr Val Gly Asn Val Ile Gly Asn Tyr His  
65 70 75 80

Pro His Gly Asp Thr Ser Val Tyr Asp Ala Met Val Arg Leu Ser Gln  
85 90 95

Pro Trp Lys Met Arg His Pro Leu Val Asp Met His Gly Asn Lys Gly  
100 105 110

Ser Met Asp Gly Asp Pro Pro Ala Ala Met Arg Tyr Thr Glu Ala Arg  
115 120 125

Leu Ser Lys Ile Ala Ser Asp Leu Leu Ala Asp Ile Asp Lys Glu Thr  
130 135 140

Val Asp His Val Leu Asn Phe Asp Asp Thr Thr Glu Glu Pro Thr Val  
145 150 155 160

Leu Pro Ala Arg Phe Pro Asn Leu Leu Val Asn Gly Ala Ser Gly Ile  
165 170 175

Ser Ala Gly Tyr Ala Thr Asp Ile Pro Pro His Asn Leu Ser Glu Val  
180 185 190

Ile Asp Ala Thr Ile His Leu Ile Asn His Pro Asn Ala Arg Leu Glu  
195 200 205

Thr Leu Met Asp Tyr Ile Gln Gly Pro Asp Phe Pro Thr Gly Gly Ile  
210 215 220

Ile Gln Gly Lys Ser Gly Leu Lys Lys Ala Tyr Gln Thr Gly Lys Gly  
225 230 235 240

Lys Ile Ile Ile Arg Ala Lys Ala Asp Ile Glu Ala Ile Arg Gly Gly  
245 250 255

Lys Ser Gln Ile Val Ile Ser Gln Ile Pro Tyr Glu Val Asn Lys Ala  
260 265 270

Arg Leu Val Gln Lys Ile Asp Asp Ile Arg Ile Asn Lys Lys Ile Asp  
275 280 285

Gly Ile Ala Asp Val Arg Asp Glu Ser Asp Arg Ser Gly Leu Arg Ile  
290 295 300

Val Val Glu Thr Lys Lys Asp Gly Asp Gly Glu Gly Ile Leu Thr Tyr  
305 310 315 320

Leu Leu Lys Asn Thr Asp Leu Gln Val Thr Tyr Asn Leu Asn Met Val  
325 330 335

Ala Ile Asp Lys Lys Arg Pro Gln Gln Val Ser Leu Lys Gln Ile Leu  
340 345 350

Ser Ser Tyr Leu Asp His Lys Arg Thr Val Val Gln Asn Arg Thr Arg  
355 360 365

Tyr Leu Leu Ala Lys Ala Lys Asp Arg Gln His Ile Val Gln Gly Leu  
370 375 380

Ile Lys Ala Ile Ser Ile Leu Asp Asp Leu Ile Gln Thr Ile Arg Ala  
385 390 395 400

Ser Glu Asn Lys Ala Asn Ala Lys Glu Asn Ile Ile Gln Ala Tyr Gly  
405 410 415

Phe Ser Gln Asp Gln Ala Glu Ala Ile Val Ser Leu Gln Leu Tyr Arg  
420 425 430

Leu Thr Asn Thr Asp Ile Lys Asp Leu Gln Ala Glu Ala Lys Asp Leu  
435 440 445

Ala Gln Ala Ile Leu Thr Tyr Gln Asp Leu Leu Thr Asn Lys Ala Ser  
450 455 460

Leu Asp Ala Leu Met Lys Glu Glu Leu Lys Glu Val Lys Gln Ala Tyr  
465 470 475 480

Gly Glu Asp Arg Leu Thr Gln Val Gln Asp Lys Ile Glu Lys Leu Glu  
485 490 495

Ile Glu Thr Gln Val Leu Val Ser Glu Glu Asp Val Met Val Thr Val  
500 505 510

Thr Gln Gly Gly Tyr Leu Lys Arg Thr Ser Ile Arg Ser Tyr Lys Ala  
515 520 525

Ser Gln Val Glu Glu Leu Gly Arg Arg Glu Asp Asp Leu Val Ile Phe  
530 535 540

Met Gln Glu Leu Ser Thr Leu Asp Gln Leu Leu Ile Phe Thr Ser Lys  
545 550 555 560

Gly Asn Val Val Asn Arg Pro Val His Glu Leu Pro Asp Ile Lys Trp  
565 570 575

Lys Asp Ile Gly Glu His Leu Ser Arg Thr Ile Pro Leu Gly Glu Asp  
580 585 590

Glu Glu Leu Ile Lys Val Tyr Pro Tyr Arg Glu Leu Asp Ala Gly Lys  
595 600 605

Arg Tyr Val Phe Ile Thr Arg Asp Gly Tyr Ile Lys Gln Ser Pro Glu  
610 615 620

Thr Glu Phe Glu Pro Lys Arg Thr Tyr Lys Ser Arg Ala Ser Thr Ala  
625 630 635 640

Ile Lys Leu Lys Ser Asp Gln Asp Arg Leu Gln Ala Val Tyr Tyr Ile  
645 650 655

Pro Asp Gln Glu Asp Tyr Asp Val Phe Leu Ala Ser Tyr Lys Gly Tyr  
660 665 670

Gly Leu Lys Tyr Gly Leu Glu Glu Val Ser Glu Val Gly Ala Gln Ala  
675 680 685

Ala Gly Val Lys Ser Met Asn Leu Lys Glu Gly Asp His Val Gln Asp  
690 695 700

Gly Leu Val Phe Lys Arg Lys Gln Phe Gln Glu Ala Leu Phe Ile Thr



705                      710                      715                      720

Gln Arg Ala Ser Val Lys Lys Met Ala Leu His Asp Phe Asp Arg Thr  
                          725                      730                      735

Ser Arg Ala Lys Arg Gly Leu Gln Ile Leu Arg Glu Leu Lys Arg Asn  
                          740                      745                      750

Pro His Arg Ile Gln Phe Met Ile Gly Ile Ser Gln Asn Lys Phe Leu  
                          755                      760                      765

Val Asn Leu Leu Thr Asp Thr Lys Lys Leu Val Gln Ile Asn Pro Asp  
                          770                      775                      780

Asp Tyr Thr Val Ser Asn Arg His Asn Asn Gly Ser Phe Val Leu Asp  
 785                      790                      795                      800

Thr Ser Arg Asp Gly Lys Pro Val Ser Tyr Tyr Leu Ser Asp Asn Asp  
                          805                      810                      815

Ser His Leu

<210> 61  
 <211> 1113  
 <212> DNA  
 <213> *Alloioococcus otitidis*

<220>  
 <221> CDS  
 <222> (4) .. (1113)  
 <223>

<400> 61

tta	gtg	gtt	gag	aca	aaa	tca	aaa	cta	gaa	aat	gca	gta	aac	acc	ctc	48
	Met	Val	Glu	Thr	Lys	Ser	Lys	Leu	Glu	Asn	Ala	Val	Asn	Thr	Leu	
1					5				10					15		
att	aaa	gac	ttg	aaa	aat	aaa	aaa	gag	tcg	acc	att	tct	tat	att	gac	96
Ile	Lys	Asp	Leu	Lys	Asn	Lys	Lys	Glu	Ser	Thr	Ile	Ser	Tyr	Ile	Asp	
			20					25						30		
ctc	agc	aac	aaa	att	gct	gaa	ccc	ttc	gaa	ctt	gaa	agt	gaa	gcc	atg	144
Leu	Ser	Asn	Lys	Ile	Ala	Glu	Pro	Phe	Glu	Leu	Glu	Ser	Glu	Ala	Met	
			35					40						45		
gac	aag	tta	atc	cag	caa	tta	gaa	gat	gat	ggg	att	ggt	gta	gtt	gac	192
Asp	Lys	Leu	Ile	Gln	Gln	Leu	Glu	Asp	Asp	Gly	Ile	Gly	Val	Val	Asp	
		50					55					60				

caa gac ggt aat ccc ttg gcc aag caa cta gcc aag cag gaa gaa gaa Gln Asp Gly Asn Pro Leu Ala Lys Gln Leu Ala Lys Gln Glu Glu Glu 65 70 75	240
gca gaa aaa gcc aag gat gaa gaa atg ata gcc cca cct ggg gtt aaa Ala Glu Lys Ala Lys Asp Glu Glu Met Ile Ala Pro Pro Gly Val Lys 80 85 90 95	288
att aac gac cct gtc cgg atg tac cta aaa gaa att ggc cgg gta gat Ile Asn Asp Pro Val Arg Met Tyr Leu Lys Glu Ile Gly Arg Val Asp 100 105 110	336
ctt tta gat gct gaa gaa gaa gtg gcc cta gcc aag cgg att gaa gaa Leu Leu Asp Ala Glu Glu Glu Val Ala Leu Ala Lys Arg Ile Glu Glu 115 120 125	384
ggc gat gaa atc gct aaa caa gaa cta gct gag gct aac ttg aga ctg Gly Asp Glu Ile Ala Lys Gln Glu Leu Ala Glu Ala Asn Leu Arg Leu 130 135 140	432
gtt gtc tct att gct aaa cgg tac gtt ggc cgg ggc atg agc ttt ttg Val Val Ser Ile Ala Lys Arg Tyr Val Gly Arg Gly Met Ser Phe Leu 145 150 155	480
gac ttg atc cag gaa ggg aat atg ggc cta atg aag gca gtt gaa aaa Asp Leu Ile Gln Glu Gly Asn Met Gly Leu Met Lys Ala Val Glu Lys 160 165 170 175	528
ttt gac tac gaa aaa ggt ttc aaa ttt tca acc tat gcc acc tgg tgg Phe Asp Tyr Glu Lys Gly Phe Lys Phe Ser Thr Tyr Ala Thr Trp Trp 180 185 190	576
atc cgt caa gcc atc act cgg gcc att gcc gac caa gcc cga acc atc Ile Arg Gln Ala Ile Thr Arg Ala Ile Ala Asp Gln Ala Arg Thr Ile 195 200 205	624
cgg att ccg gtc cac atg gtc gaa act att aac aag ctg gtc cga atc Arg Ile Pro Val His Met Val Glu Thr Ile Asn Lys Leu Val Arg Ile 210 215 220	672
cag cgg cag ctc cta caa gaa cta ggc cgg gaa cca acc cca gaa gaa Gln Arg Gln Leu Leu Gln Glu Leu Gly Arg Glu Pro Thr Pro Glu Glu 225 230 235	720
att ggg gca gag atg gat ttg cca acc gaa aaa gtc aga gat att ttg Ile Gly Ala Glu Met Asp Leu Pro Thr Glu Lys Val Arg Asp Ile Leu 240 245 250 255	768
aaa att tcc caa gaa ccc gtc tcc ctt gaa acc cca att ggg gaa gaa Lys Ile Ser Gln Glu Pro Val Ser Leu Glu Thr Pro Ile Gly Glu Glu 260 265 270	816
gaa gat tcc cac ctg gga gac ttt att gaa gat gat ggg gcc ttg tcg Glu Asp Ser His Leu Gly Asp Phe Ile Glu Asp Asp Gly Ala Leu Ser 275 280 285	864
cca tct gat aat gca gct tat gag ctg ttg aaa ggg gaa ctc aaa gga	912

Pro Ser Asp Asn Ala Ala Tyr Glu Leu Leu Lys Gly Glu Leu Lys Gly  
 290 295 300

gtc tta gac acc cta act gac cgg gaa gaa aat gtc ttg cgc ctc cgt 960  
 Val Leu Asp Thr Leu Thr Asp Arg Glu Glu Asn Val Leu Arg Leu Arg  
 305 310 315

ttt ggc cta gat gat ggc cgt caa cgt act tta gaa gat gtc ggt aag 1008  
 Phe Gly Leu Asp Asp Gly Arg Gln Arg Thr Leu Glu Asp Val Gly Lys  
 320 325 330 335

gtc ttt ggg gtc acc cgg gag cgg atc cgt caa att gaa gcg aag gcc 1056  
 Val Phe Gly Val Thr Arg Glu Arg Ile Arg Gln Ile Glu Ala Lys Ala  
 340 345 350

ctc cgc aaa ctc cgc cac cct agc cgg tcc aaa caa tta aaa gac ttt 1104  
 Leu Arg Lys Leu Arg His Pro Ser Arg Ser Lys Gln Leu Lys Asp Phe  
 355 360 365

tta gaa tag 1113  
 Leu Glu

<210> 62  
 <211> 369  
 <212> PRT  
 <213> *Alloiococcus otitidis*

<400> 62  
 Met Val Glu Thr Lys Ser Lys Leu Glu Asn Ala Val Asn Thr Leu Ile  
 1 5 10 15

Lys Asp Leu Lys Asn Lys Lys Glu Ser Thr Ile Ser Tyr Ile Asp Leu  
 20 25 30

Ser Asn Lys Ile Ala Glu Pro Phe Glu Leu Glu Ser Glu Ala Met Asp  
 35 40 45

Lys Leu Ile Gln Gln Leu Glu Asp Asp Gly Ile Gly Val Val Asp Gln  
 50 55 60

Asp Gly Asn Pro Leu Ala Lys Gln Leu Ala Lys Gln Glu Glu Glu Ala  
 65 70 75 80

Glu Lys Ala Lys Asp Glu Glu Met Ile Ala Pro Pro Gly Val Lys Ile  
 85 90 95

Asn Asp Pro Val Arg Met Tyr Leu Lys Glu Ile Gly Arg Val Asp Leu  
 100 105 110

Leu Asp Ala Glu Glu Glu Val Ala Leu Ala Lys Arg Ile Glu Glu Gly  
115 120 125

Asp Glu Ile Ala Lys Gln Glu Leu Ala Glu Ala Asn Leu Arg Leu Val  
130 135 140

Val Ser Ile Ala Lys Arg Tyr Val Gly Arg Gly Met Ser Phe Leu Asp  
145 150 155 160

Leu Ile Gln Glu Gly Asn Met Gly Leu Met Lys Ala Val Glu Lys Phe  
165 170 175

Asp Tyr Glu Lys Gly Phe Lys Phe Ser Thr Tyr Ala Thr Trp Trp Ile  
180 185 190

Arg Gln Ala Ile Thr Arg Ala Ile Ala Asp Gln Ala Arg Thr Ile Arg  
195 200 205

Ile Pro Val His Met Val Glu Thr Ile Asn Lys Leu Val Arg Ile Gln  
210 215 220

Arg Gln Leu Leu Gln Glu Leu Gly Arg Glu Pro Thr Pro Glu Glu Ile  
225 230 235 240

Gly Ala Glu Met Asp Leu Pro Thr Glu Lys Val Arg Asp Ile Leu Lys  
245 250 255

Ile Ser Gln Glu Pro Val Ser Leu Glu Thr Pro Ile Gly Glu Glu Glu  
260 265 270

Asp Ser His Leu Gly Asp Phe Ile Glu Asp Asp Gly Ala Leu Ser Pro  
275 280 285

Ser Asp Asn Ala Ala Tyr Glu Leu Leu Lys Gly Glu Leu Lys Gly Val  
290 295 300

Leu Asp Thr Leu Thr Asp Arg Glu Glu Asn Val Leu Arg Leu Arg Phe  
305 310 315 320

Gly Leu Asp Asp Gly Arg Gln Arg Thr Leu Glu Asp Val Gly Lys Val  
325 330 335

Phe Gly Val Thr Arg Glu Arg Ile Arg Gln Ile Glu Ala Lys Ala Leu

340

345

350

Arg Lys Leu Arg His Pro Ser Arg Ser Lys Gln Leu Lys Asp Phe Leu  
 355 360 365

Glu

&lt;210&gt; 63

&lt;211&gt; 1854

&lt;212&gt; DNA

<213> *Alloioococcus otitidis*

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(1854)

&lt;223&gt;

&lt;400&gt; 63

atg gtt aga ata cct gaa gag acc att aat caa ata cga agc cag gca 48  
 Met Val Arg Ile Pro Glu Glu Thr Ile Asn Gln Ile Arg Ser Gln Ala  
 1 5 10 15

gat att gtc gat gtc att ggc caa tac ttg gac tta aac aag tct ggg 96  
 Asp Ile Val Asp Val Ile Gly Gln Tyr Leu Asp Leu Asn Lys Ser Gly  
 20 25 30

gcc aat tac ttt gcc cac tgc ccc ttc cat gaa gac agc acg cct tct 144  
 Ala Asn Tyr Phe Ala His Cys Pro Phe His Glu Asp Ser Thr Pro Ser  
 35 40 45

ttt tcg gtc aac aga gac aag caa att tat aag tgc ttt tct tgc aaa 192  
 Phe Ser Val Asn Arg Asp Lys Gln Ile Tyr Lys Cys Phe Ser Cys Lys  
 50 55 60

cga ggt ggc agt gtc ttt agc ttt ata caa gag aag gag gga ctt tcc 240  
 Arg Gly Gly Ser Val Phe Ser Phe Ile Gln Glu Lys Glu Gly Leu Ser  
 65 70 75 80

ttc cca gaa tcg gtt ctt aaa gtg gca gac tta gct aat gtg gac ctt 288  
 Phe Pro Glu Ser Val Leu Lys Val Ala Asp Leu Ala Asn Val Asp Leu  
 85 90 95

gat ccg gcc tta aaa gaa gct gtc caa ggc caa cct gac aaa gcc gat 336  
 Asp Pro Ala Leu Lys Glu Ala Val Gln Gly Gln Pro Asp Lys Ala Asp  
 100 105 110

tct ccc tac cga gac ctc tat acc atc cat gac cag gcc aag gac tac 384  
 Ser Pro Tyr Arg Asp Leu Tyr Thr Ile His Asp Gln Ala Lys Asp Tyr  
 115 120 125

tac cag tat atc ctc tta aag gcc cag gtg gga gaa gtt gct tac gac 432  
 Tyr Gln Tyr Ile Leu Leu Lys Ala Gln Val Gly Glu Val Ala Tyr Asp  
 130 135 140

tat ctc cag aat cgt ggg att tcc aga gag gtg atg gaa gag ttc gaa Tyr Leu Gln Asn Arg Gly Ile Ser Arg Glu Val Met Glu Glu Phe Glu 145 150 155 160	480
ctg ggt tat tct ccc agc caa agg gag tcg ctc cac ctt tat ttg cag Leu Gly Tyr Ser Pro Ser Gln Arg Glu Ser Leu His Leu Tyr Leu Gln 165 170 175	528
tcc caa gac cag gcg gac ttg aca gat gac tta ctg gaa gaa acc ggc Ser Gln Asp Gln Ala Asp Leu Thr Asp Asp Leu Leu Glu Glu Thr Gly 180 185 190	576
ctt ttt tcc aaa aga gaa gtg gaa agt gat agt ttt aaa gac cgc ttt Leu Phe Ser Lys Arg Glu Val Glu Ser Asp Ser Phe Lys Asp Arg Phe 195 200 205	624
gcc aag cgg atc atc ttc ccc tta aag aac tta caa ggg cag acg gtg Ala Lys Arg Ile Ile Phe Pro Leu Lys Asn Leu Gln Gly Gln Thr Val 210 215 220	672
ggc ttt tcg ggc cgg tat ttc caa gat gag cct aac cag gac ttc cat Gly Phe Ser Gly Arg Tyr Phe Gln Asp Glu Pro Asn Gln Asp Phe His 225 230 235 240	720
cat gcc aag tat tta aac agt cca gaa acc aaa ata ttc aat aaa cgg His Ala Lys Tyr Leu Asn Ser Pro Glu Thr Lys Ile Phe Asn Lys Arg 245 250 255	768
cgg acc ctc ttt aac tac cac cag gcc aag gcc tac att cgt cgg gcc Arg Thr Leu Phe Asn Tyr His Gln Ala Lys Ala Tyr Ile Arg Arg Ala 260 265 270	816
aag gaa gtt gtc tta ttc gaa ggt tac atg gat gtg att gct gct tgg Lys Glu Val Val Leu Phe Glu Gly Tyr Met Asp Val Ile Ala Ala Trp 275 280 285	864
caa gcg ggg gtc aaa aat ggc tta gct tcc atg ggg acc agt ata aca Gln Ala Gly Val Lys Asn Gly Leu Ala Ser Met Gly Thr Ser Ile Thr 290 295 300	912
gct gac caa gtc cag acc atg caa agg att gct gac acc tta gtc ttg Ala Asp Gln Val Gln Thr Met Gln Arg Ile Ala Asp Thr Leu Val Leu 305 310 315 320	960
gcc ttt gac ggg gat gaa gct ggc ctt gaa tcc agc aaa aag atc ctg Ala Phe Asp Gly Asp Glu Ala Gly Leu Glu Ser Ser Lys Lys Ile Leu 325 330 335	1008
gat gac tta agc ttg acc agc aag ctt caa att gaa gtg gtc att ttc Asp Asp Leu Ser Leu Thr Ser Lys Leu Gln Ile Glu Val Val Ile Phe 340 345 350	1056
cct aaa aaa atg gac ccg gat gaa tat att aga gaa aat gga cca gaa Pro Lys Lys Met Asp Pro Asp Glu Tyr Ile Arg Glu Asn Gly Pro Glu 355 360 365	1104
gcc ttt caa aat ctc atc caa cat ggt agg atg act gtc tac caa ttc	1152

Ala Phe Gln Asn Leu Ile Gln His Gly Arg Met Thr Val Tyr Gln Phe	
370 375 380	
tta aaa gaa tac ttt aaa aaa tcc tac aat cta gat aac gac tcg gac	1200
Leu Lys Glu Tyr Phe Lys Lys Ser Tyr Asn Leu Asp Asn Asp Ser Asp	
385 390 395 400	
cgg ttg aaa ttt atc caa acc atg acc aat aaa att ggc aag cta gct	1248
Arg Leu Lys Phe Ile Gln Thr Met Thr Asn Lys Ile Gly Lys Leu Ala	
405 410 415	
tcc ccc ttg gaa agg gaa gtc tat gcc aag gat ttg gca gaa gaa ttt	1296
Ser Pro Leu Glu Arg Glu Val Tyr Ala Lys Asp Leu Ala Glu Glu Phe	
420 425 430	
aac ctg tct tat gat acg att ata agc caa gtt caa agt gaa gcc act	1344
Asn Leu Ser Tyr Asp Thr Ile Ile Ser Gln Val Gln Ser Glu Ala Thr	
435 440 445	
cta aac cag caa gag gct ttg aaa aag gac cgg cat aag gaa ttt tct	1392
Leu Asn Gln Gln Glu Ala Leu Lys Lys Asp Arg His Lys Glu Phe Ser	
450 455 460	
caa gca aga gtg gaa gtc aaa gcc cca agt agt caa aag act aag att	1440
Gln Ala Arg Val Glu Val Lys Ala Pro Ser Ser Gln Lys Thr Lys Ile	
465 470 475 480	
gac cgg gcc cag gaa aaa ctt tta aac cga ctc ttt tac tat ccc caa	1488
Asp Arg Ala Gln Glu Lys Leu Leu Asn Arg Leu Phe Tyr Tyr Pro Gln	
485 490 495	
gtt caa gag atc atc gat gct tat aat ccg gac ttt gaa ttt aaa acg	1536
Val Gln Glu Ile Ile Asp Ala Tyr Asn Pro Asp Phe Glu Phe Lys Thr	
500 505 510	
gaa gtc cac cag cgg att tac ctc ttg ttt tta gaa tac agc cag gaa	1584
Glu Val His Gln Arg Ile Tyr Leu Leu Phe Leu Glu Tyr Ser Gln Glu	
515 520 525	
aat gat agc att gat tct ttc atc gat ttt gtc aaa gac aag gag acg	1632
Asn Asp Ser Ile Asp Ser Phe Ile Asp Phe Val Lys Asp Lys Glu Thr	
530 535 540	
aaa gag gtc ata tct gat ata atg tgg aca tcc att gag gtc gaa ccc	1680
Lys Glu Val Ile Ser Asp Ile Met Trp Thr Ser Ile Glu Val Glu Pro	
545 550 555 560	
tca gat gaa gaa atc cta gac tac ttg gac tac att gac caa acc tac	1728
Ser Asp Glu Glu Ile Leu Asp Tyr Leu Asp Tyr Ile Asp Gln Thr Tyr	
565 570 575	
ccc ctg gag caa aaa cgc caa gac tgc ttg gag gaa gtc aaa gca gct	1776
Pro Leu Glu Gln Lys Arg Gln Asp Cys Leu Glu Glu Val Lys Ala Ala	
580 585 590	
aaa cag tcc ggt aat aag aag cga gag ctg gaa tta acc aat caa tta	1824
Lys Gln Ser Gly Asn Lys Lys Arg Glu Leu Glu Leu Thr Asn Gln Leu	

595

600

605

att gaa ata aac cgt atg cta aaa caa taa  
 ile Glu ile Asn Arg Met Leu Lys Gln  
 610 615

1854

&lt;210&gt; 64

&lt;211&gt; 617

&lt;212&gt; PRT

&lt;213&gt; Alloiococcus otitidis

&lt;400&gt; 64

Met Val Arg ile Pro Glu Glu Thr ile Asn Gln ile Arg Ser Gln Ala  
 1 5 10 15

Asp ile Val Asp Val ile Gly Gln Tyr Leu Asp Leu Asn Lys Ser Gly  
 20 25 30

Ala Asn Tyr Phe Ala His Cys Pro Phe His Glu Asp Ser Thr Pro Ser  
 35 40 45

Phe Ser Val Asn Arg Asp Lys Gln ile Tyr Lys Cys Phe Ser Cys Lys  
 50 55 60

Arg Gly Gly Ser Val Phe Ser Phe ile Gln Glu Lys Glu Gly Leu Ser  
 65 70 75 80

Phe Pro Glu Ser Val Leu Lys Val Ala Asp Leu Ala Asn Val Asp Leu  
 85 90 95

Asp Pro Ala Leu Lys Glu Ala Val Gln Gly Gln Pro Asp Lys Ala Asp  
 100 105 110

Ser Pro Tyr Arg Asp Leu Tyr Thr ile His Asp Gln Ala Lys Asp Tyr  
 115 120 125

Tyr Gln Tyr ile Leu Leu Lys Ala Gln Val Gly Glu Val Ala Tyr Asp  
 130 135 140

Tyr Leu Gln Asn Arg Gly ile Ser Arg Glu Val Met Glu Glu Phe Glu  
 145 150 155 160

Leu Gly Tyr Ser Pro Ser Gln Arg Glu Ser Leu His Leu Tyr Leu Gln  
 165 170 175



Ser Gln Asp Gln Ala Asp Leu Thr Asp Asp Leu Leu Glu Glu Thr Gly  
180 185 190

Leu Phe Ser Lys Arg Glu Val Glu Ser Asp Ser Phe Lys Asp Arg Phe  
195 200 205

Ala Lys Arg Ile Ile Phe Pro Leu Lys Asn Leu Gln Gly Gln Thr Val  
210 215 220

Gly Phe Ser Gly Arg Tyr Phe Gln Asp Glu Pro Asn Gln Asp Phe His  
225 230 235 240

His Ala Lys Tyr Leu Asn Ser Pro Glu Thr Lys Ile Phe Asn Lys Arg  
245 250 255

Arg Thr Leu Phe Asn Tyr His Gln Ala Lys Ala Tyr Ile Arg Arg Ala  
260 265 270

Lys Glu Val Val Leu Phe Glu Gly Tyr Met Asp Val Ile Ala Ala Trp  
275 280 285

Gln Ala Gly Val Lys Asn Gly Leu Ala Ser Met Gly Thr Ser Ile Thr  
290 295 300

Ala Asp Gln Val Gln Thr Met Gln Arg Ile Ala Asp Thr Leu Val Leu  
305 310 315 320

Ala Phe Asp Gly Asp Glu Ala Gly Leu Glu Ser Ser Lys Lys Ile Leu  
325 330 335

Asp Asp Leu Ser Leu Thr Ser Lys Leu Gln Ile Glu Val Val Ile Phe  
340 345 350

Pro Lys Lys Met Asp Pro Asp Glu Tyr Ile Arg Glu Asn Gly Pro Glu  
355 360 365

Ala Phe Gln Asn Leu Ile Gln His Gly Arg Met Thr Val Tyr Gln Phe  
370 375 380

Leu Lys Glu Tyr Phe Lys Lys Ser Tyr Asn Leu Asp Asn Asp Ser Asp  
385 390 395 400

Arg Leu Lys Phe Ile Gln Thr Met Thr Asn Lys Ile Gly Lys Leu Ala

405

410

415

Ser Pro Leu Glu Arg Glu Val Tyr Ala Lys Asp Leu Ala Glu Glu Phe  
420 425 430

Asn Leu Ser Tyr Asp Thr Ile Ile Ser Gln Val Gln Ser Glu Ala Thr  
435 440 445

Leu Asn Gln Gln Glu Ala Leu Lys Lys Asp Arg His Lys Glu Phe Ser  
450 455 460

Gln Ala Arg Val Glu Val Lys Ala Pro Ser Ser Gln Lys Thr Lys Ile  
465 470 475 480

Asp Arg Ala Gln Glu Lys Leu Leu Asn Arg Leu Phe Tyr Tyr Pro Gln  
485 490 495

Val Gln Glu Ile Ile Asp Ala Tyr Asn Pro Asp Phe Glu Phe Lys Thr  
500 505 510

Glu Val His Gln Arg Ile Tyr Leu Leu Phe Leu Glu Tyr Ser Gln Glu  
515 520 525

Asn Asp Ser Ile Asp Ser Phe Ile Asp Phe Val Lys Asp Lys Glu Thr  
530 535 540

Lys Glu Val Ile Ser Asp Ile Met Trp Thr Ser Ile Glu Val Glu Pro  
545 550 555 560

Ser Asp Glu Glu Ile Leu Asp Tyr Leu Asp Tyr Ile Asp Gln Thr Tyr  
565 570 575

Pro Leu Glu Gln Lys Arg Gln Asp Cys Leu Glu Glu Val Lys Ala Ala  
580 585 590

Lys Gln Ser Gly Asn Lys Lys Arg Glu Leu Glu Leu Thr Asn Gln Leu  
595 600 605

Ile Glu Ile Asn Arg Met Leu Lys Gln  
610 615

<210> 65  
<211> 987

&lt;212&gt; DNA

<213> *Alloiococcus otitidis*

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (55)..(987)

&lt;223&gt;

&lt;400&gt; 65

```

gccagacaat cggacaacta ttaccggaca actttaagcc aggggacatg gaat atg      57
                                         Met
                                         1

gaa aac aat gaa aac aat gaa aac aaa gat agc aaa aca ttt aaa tca      105
Glu Asn Asn Glu Asn Asn Glu Asn Lys Asp Ser Lys Thr Phe Lys Ser
      5              10              15

ggt ttt gtc acc ctt ctt ggc cgg ccc aat gtg ggc aag tca acc ctg      153
Gly Phe Val Thr Leu Leu Gly Arg Pro Asn Val Gly Lys Ser Thr Leu
      20              25              30

ctc aac caa ata tta ggc cag aag att acc att atc agt gac aaa ccc      201
Leu Asn Gln Ile Leu Gly Gln Lys Ile Thr Ile Ile Ser Asp Lys Pro
      35              40              45

caa aca acc cgg aat aaa atc cag ggt att tac acc gac caa gcg ggg      249
Gln Thr Thr Arg Asn Lys Ile Gln Gly Ile Tyr Thr Asp Gln Ala Gly
      50              55              60              65

caa att gtc ttt atc gac aca cct ggt ata cat aaa ccc aag cac cgc      297
Gln Ile Val Phe Ile Asp Thr Pro Gly Ile His Lys Pro Lys His Arg
      70              75              80

ctg ggc cgg ttt atg gtg gat tcg gct atg atg tcg acc atc aat gag gtg      345
Leu Gly Arg Phe Met Val Asp Ser Ala Met Ser Thr Ile Asn Glu Val
      85              90              95

gac ctg gtc tta ttt gtg gtc aat gtc agg gaa aag att ggc ccg ggg      393
Asp Leu Val Leu Phe Val Val Asn Val Arg Glu Lys Ile Gly Pro Gly
      100             105             110

gac cgg ttc att atc gac aag ttg cga acc atc gat acg cca gtt ttt      441
Asp Arg Phe Ile Ile Asp Lys Leu Arg Thr Ile Asp Thr Pro Val Phe
      115             120             125

tta att att aac cag att gac cag gtc gat cca aca gac ctc cta ccg      489
Leu Ile Ile Asn Gln Ile Asp Gln Val Asp Pro Thr Asp Leu Leu Pro
      130             135             140             145

gtt att agc gac tac caa gag gaa ttc gac ttt gcc gaa gtg gtt cca      537
Val Ile Ser Asp Tyr Gln Glu Glu Phe Asp Phe Ala Glu Val Val Pro
      150             155             160

act tca ggc ttg gaa ggg gaa aat atc cag gag ctc att caa acc atc      585
Thr Ser Gly Leu Glu Gly Glu Asn Ile Gln Glu Leu Ile Gln Thr Ile
      165             170             175

```

aag tct tac cta cct gtt gga ccc caa ttt tac ccg gac gac cag gtc 633  
Lys Ser Tyr Leu Pro Val Gly Pro Gln Phe Tyr Pro Asp Asp Gln Val  
180 185 190  
  
tcg gac cac ccc gaa tac ttt att att tca gaa ctc atc cgg gag aag 681  
Ser Asp His Pro Glu Tyr Phe Ile Ile Ser Glu Leu Ile Arg Glu Lys  
195 200 205  
  
gtt tta gac ttg gct aga gaa gag att cct cat tca gta gca gta gta 729  
Val Leu Asp Leu Ala Arg Glu Glu Ile Pro His Ser Val Ala Val Val  
210 215 220 225  
  
act gag aag gta gac cga aac caa gat ggt aaa gtc caa acc tat gcc 777  
Thr Glu Lys Val Asp Arg Asn Gln Asp Gly Lys Val Gln Thr Tyr Ala  
230 235 240  
  
acc att att gtc gaa cgc aag agc caa aag ggg att att atc ggc aag 825  
Thr Ile Ile Val Glu Arg Lys Ser Gln Lys Gly Ile Ile Ile Gly Lys  
245 250 255  
  
caa ggg tcc atg att aaa aaa att ggt agc cta gct cgg cga gat att 873  
Gln Gly Ser Met Ile Lys Lys Ile Gly Ser Leu Ala Arg Arg Asp Ile  
260 265 270  
  
gag aaa cta ctg gga gat aag att tac ttg gaa ctc tgg gtt aaa gtc 921  
Glu Lys Leu Leu Gly Asp Lys Ile Tyr Leu Glu Leu Trp Val Lys Val  
275 280 285  
  
caa aga gac tgg cgg gac aag ccc agt cgc tta gaa gac ttt ggc tac 969  
Gln Arg Asp Trp Arg Asp Lys Pro Ser Arg Leu Glu Asp Phe Gly Tyr  
290 295 300 305  
  
aat gaa gac aac tat tag 987  
Asn Glu Asp Asn Tyr  
310

&lt;210&gt; 66

&lt;211&gt; 310

&lt;212&gt; PRT

&lt;213&gt; Alloiococcus otitidis

&lt;400&gt; 66

Met Glu Asn Asn Glu Asn Asn Glu Asn Lys Asp Ser Lys Thr Phe Lys  
1 5 10 15

Ser Gly Phe Val Thr Leu Leu Gly Arg Pro Asn Val Gly Lys Ser Thr  
20 25 30

Leu Leu Asn Gln Ile Leu Gly Gln Lys Ile Thr Ile Ile Ser Asp Lys  
35 40 45

Pro Gln Thr Thr Arg Asn Lys Ile Gln Gly Ile Tyr Thr Asp Gln Ala  
50 55 60

Gly Gln Ile Val Phe Ile Asp Thr Pro Gly Ile His Lys Pro Lys His  
 65 70 75 80

Arg Leu Gly Arg Phe Met Val Asp Ser Ala Met Ser Thr Ile Asn Glu  
 85 90 95

Val Asp Leu Val Leu Phe Val Val Asn Val Arg Glu Lys Ile Gly Pro  
 100 105 110

Gly Asp Arg Phe Ile Ile Asp Lys Leu Arg Thr Ile Asp Thr Pro Val  
 115 120 125

Phe Leu Ile Ile Asn Gln Ile Asp Gln Val Asp Pro Thr Asp Leu Leu  
 130 135 140

Pro Val Ile Ser Asp Tyr Gln Glu Glu Phe Asp Phe Ala Glu Val Val  
 145 150 155 160

Pro Thr Ser Gly Leu Glu Gly Glu Asn Ile Gln Glu Leu Ile Gln Thr  
 165 170 175

Ile Lys Ser Tyr Leu Pro Val Gly Pro Gln Phe Tyr Pro Asp Asp Gln  
 180 185 190

Val Ser Asp His Pro Glu Tyr Phe Ile Ile Ser Glu Leu Ile Arg Glu  
 195 200 205

Lys Val Leu Asp Leu Ala Arg Glu Glu Ile Pro His Ser Val Ala Val  
 210 215 220

Val Thr Glu Lys Val Asp Arg Asn Gln Asp Gly Lys Val Gln Thr Tyr  
 225 230 235 240

Ala Thr Ile Ile Val Glu Arg Lys Ser Gln Lys Gly Ile Ile Ile Gly  
 245 250 255

Lys Gln Gly Ser Met Ile Lys Lys Ile Gly Ser Leu Ala Arg Arg Asp  
 260 265 270

Ile Glu Lys Leu Leu Gly Asp Lys Ile Tyr Leu Glu Leu Trp Val Lys  
 275 280 285

Val Gln Arg Asp Trp Arg Asp Lys Pro Ser Arg Leu Glu Asp Phe Gly  
 290 295 300

Tyr Asn Glu Asp Asn Tyr  
 305 310

<210> 67  
 <211> 1557  
 <212> DNA  
 <213> *Alloiococcus otitidis*

<220>  
 <221> CDS  
 <222> (46)..(1557)  
 <223>

<400> 67  
 catgtctatg ttactttaat cagtggaaaa caagaggaga tcatt gtg att tcc tct 57  
 Met Ile Ser Ser  
 1

ttc tat tta gta gga gtc ttg aga ttg agt agt gaa aat aaa tta acc 105  
 Phe Tyr Leu Val Gly Val Leu Arg Leu Ser Ser Glu Asn Lys Leu Thr  
 5 10 15 20

ttc aaa cac ttc ctt gca aac cag ttg acc aaa cga gac aat tta caa 153  
 Phe Lys His Phe Leu Ala Asn Gln Leu Thr Lys Arg Asp Asn Leu Gln  
 25 30 35

atc ccc cgt tgg caa att ttt gcc gtt tta ttt aca gga gcc gtg att 201  
 Ile Pro Arg Trp Gln Ile Phe Ala Val Leu Phe Thr Gly Ala Val Ile  
 40 45 50

gtg gtt ctc aac caa acg gcc atg tct acc gcc ttg cct aat atg att 249  
 Val Val Leu Asn Gln Thr Ala Met Ser Thr Ala Leu Pro Asn Met Ile  
 55 60 65

gaa agt ttg ggc att gac cct agc cta ggc cag tgg att gtc tcg ggt 297  
 Glu Ser Leu Gly Ile Asp Pro Ser Leu Gly Gln Trp Ile Val Ser Gly  
 70 75 80

tat acc ttg gtc aaa ggg att atg gtc ccc ata acc gcc ttt gcc atg 345  
 Tyr Thr Leu Val Lys Gly Ile Met Val Pro Ile Thr Ala Phe Ala Met  
 85 90 95 100

acc aag tac cgg aca cgg aac ttt ttt att tta atg ttg gcc ctc ttc 393  
 Thr Lys Tyr Arg Thr Arg Asn Phe Phe Ile Leu Met Leu Ala Leu Phe  
 105 110 115

tgt acc ggt agt ttt ttg act ggt ctg ggc ttt aat ttt ccg gtt gtg 441  
 Cys Thr Gly Ser Phe Leu Thr Gly Leu Gly Phe Asn Phe Pro Val Val  
 120 125 130

gtc atg ggg aca gtc atc cag ggt ata gcg gct ggg atg atc atc ccc 489

Val Met Gly Thr Val Ile Gln Gly Ile Ala Ala Gly Met Ile Ile Pro	
135 140 145	
ttg atg cag acc gtc ctc ttg acc ttg atg ccg gtt gaa agc cga ggc	537
Leu Met Gln Thr Val Leu Leu Thr Leu Met Pro Val Glu Ser Arg Gly	
150 155 160	
act gct atg ggg gta atg agt ggg gtt att ggt att ggt cca gca ctg	585
Thr Ala Met Gly Val Met Ser Gly Val Ile Gly Ile Gly Pro Ala Leu	
165 170 175 180	
ggg ccc ctt gtc ggt ggg gtc att gtt gat gct ttc acc tgg gaa att	633
Gly Pro Leu Val Gly Gly Val Ile Val Asp Ala Phe Thr Trp Glu Ile	
185 190 195	
tta ttc tac atc tgg gcc tta atc acc ctt tta ttg gtt cct tta act	681
Leu Phe Tyr Ile Trp Ala Leu Ile Thr Leu Leu Leu Val Pro Leu Thr	
200 205 210	
tgg ctg gtc tta ccc gat gta ttg cca aat gca gat tta acc att aat	729
Trp Leu Val Leu Pro Asp Val Leu Pro Asn Ala Asp Leu Thr Ile Asn	
215 220 225	
tgg gcc aat atc cgg gac tcc ctc att ggt ttt ggc ctc ctc ctc ttt	777
Trp Ala Asn Ile Arg Asp Ser Leu Ile Gly Phe Gly Leu Leu Leu Phe	
230 235 240	
agc ttg tca gtc ttt ggt tct tcc ggt ttt tct tcg gtc att gcc tgg	825
Ser Leu Ser Val Phe Gly Ser Ser Gly Phe Ser Ser Val Ile Ala Trp	
245 250 255 260	
gtc agc ttg ctt atc ggt tta gtc ttt gtc gcc aag ttt atc cac ttc	873
Val Ser Leu Leu Ile Gly Leu Val Phe Val Ala Lys Phe Ile His Phe	
265 270 275	
aac ctc aag gca gac caa cca atc tta aat ctt aga ctc ttt aaa aaa	921
Asn Leu Lys Ala Asp Gln Pro Ile Leu Asn Leu Arg Leu Phe Lys Lys	
280 285 290	
acc tat tac cgt cgg gct gtc ttg gta gcc acc ttg ggg att gtc att	969
Thr Tyr Tyr Arg Arg Ala Val Leu Val Ala Thr Leu Gly Ile Val Ile	
295 300 305	
att tct tgt cta tcc aac att atc cct att tat gtt caa act gtt agg	1017
Ile Ser Cys Leu Ser Asn Ile Ile Pro Ile Tyr Val Gln Thr Val Arg	
310 315 320	
ggc ttg ggg gct tcc ata gca ggc tta atc tta atg cca gct ggt atc	1065
Gly Leu Gly Ala Ser Ile Ala Gly Leu Ile Leu Met Pro Ala Gly Ile	
325 330 335 340	
atc aaa acc atc tta gct cct atc tca ggc aaa ctt tat gac aag gtt	1113
Ile Lys Thr Ile Leu Ala Pro Ile Ser Gly Lys Leu Tyr Asp Lys Val	
345 350 355	
gga gtg gct cgg att ggc ctt atc ggt ggt atc tta ctt tta gtt ggg	1161
Gly Val Ala Arg Ile Gly Leu Ile Gly Gly Ile Leu Leu Leu Val Gly	

360	365	370	
tcc tta tta cta gtt acc ctc aat gaa gct agc tcc ctt tac tta ctg Ser Leu Leu Leu Val Thr Leu Asn Glu Ala Ser Ser Leu Tyr Leu Leu 375 380 385			1209
atg att tac tac ggc atc tta tca gcc ggt ttt ggc ttg ttt aat atc Met Ile Tyr Tyr Gly Ile Leu Ser Ala Gly Phe Gly Leu Phe Asn Ile 390 395 400			1257
cct att acc act gct ggc atg aat att atg gcc aag gaa gat atg gga Pro Ile Thr Thr Ala Gly Met Asn Ile Met Ala Lys Glu Asp Met Gly 405 410 415 420			1305
cat gcg act tca gcc cgg caa acg gtc cgg caa atc tct tca agt ttt His Ala Thr Ser Ala Arg Gln Thr Val Arg Gln Ile Ser Ser Ser Phe 425 430 435			1353
gcc gtt tcc ctc tcc ttt atc atc atg acc ctg gtg act att gcc act Ala Val Ser Leu Ser Phe Ile Ile Met Thr Leu Val Thr Ile Ala Thr 440 445 450			1401
tcc ggc caa tcg gtg ggg gtt ttc caa gat ggc ggt ccg aca gac tta Ser Gly Gln Ser Val Gly Val Phe Gln Asp Gly Gly Pro Thr Asp Leu 455 460 465			1449
aat atg gca gga gtc cga ggc gcc ttt atc ttg gtg gct ata ttt tca Asn Met Ala Gly Val Arg Gly Ala Phe Ile Leu Val Ala Ile Phe Ser 470 475 480			1497
atc cta gcc atg atc ttg atc ttc ttt tta aaa gac cct aaa gaa aaa Ile Leu Ala Met Ile Leu Ile Phe Phe Leu Lys Asp Pro Lys Glu Lys 485 490 495 500			1545
cca gac caa tag Pro Asp Gln			1557

&lt;210&gt; 68

&lt;211&gt; 503

&lt;212&gt; PRT

<213> *Alloiococcus otitidis*

&lt;400&gt; 68

Met	Ile	Ser	Ser	Phe	Tyr	Leu	Val	Gly	Val	Leu	Arg	Leu	Ser	Ser	Glu
1				5					10					15	

Asn	Lys	Leu	Thr	Phe	Lys	His	Phe	Leu	Ala	Asn	Gln	Leu	Thr	Lys	Arg
			20						25					30	

Asp	Asn	Leu	Gln	Ile	Pro	Arg	Trp	Gln	Ile	Phe	Ala	Val	Leu	Phe	Thr
		35					40					45			

Gly	Ala	Val	Ile	Val	Val	Leu	Asn	Gln	Thr	Ala	Met	Ser	Thr	Ala	Leu
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----



50

55

60

Pro Asn Met Ile Glu Ser Leu Gly Ile Asp Pro Ser Leu Gly Gln Trp  
65 70 75 80

Ile Val Ser Gly Tyr Thr Leu Val Lys Gly Ile Met Val Pro Ile Thr  
85 90 95

Ala Phe Ala Met Thr Lys Tyr Arg Thr Arg Asn Phe Phe Ile Leu Met  
100 105 110

Leu Ala Leu Phe Cys Thr Gly Ser Phe Leu Thr Gly Leu Gly Phe Asn  
115 120 125

Phe Pro Val Val Val Met Gly Thr Val Ile Gln Gly Ile Ala Ala Gly  
130 135 140

Met Ile Ile Pro Leu Met Gln Thr Val Leu Leu Thr Leu Met Pro Val  
145 150 155 160

Glu Ser Arg Gly Thr Ala Met Gly Val Met Ser Gly Val Ile Gly Ile  
165 170 175

Gly Pro Ala Leu Gly Pro Leu Val Gly Gly Val Ile Val Asp Ala Phe  
180 185 190

Thr Trp Glu Ile Leu Phe Tyr Ile Trp Ala Leu Ile Thr Leu Leu Leu  
195 200 205

Val Pro Leu Thr Trp Leu Val Leu Pro Asp Val Leu Pro Asn Ala Asp  
210 215 220

Leu Thr Ile Asn Trp Ala Asn Ile Arg Asp Ser Leu Ile Gly Phe Gly  
225 230 235 240

Leu Leu Leu Phe Ser Leu Ser Val Phe Gly Ser Ser Gly Phe Ser Ser  
245 250 255

Val Ile Ala Trp Val Ser Leu Leu Ile Gly Leu Val Phe Val Ala Lys  
260 265 270

Phe Ile His Phe Asn Leu Lys Ala Asp Gln Pro Ile Leu Asn Leu Arg  
275 280 285

Leu Phe Lys Lys Thr Tyr Tyr Arg Arg Ala Val Leu Val Ala Thr Leu  
290 295 300

Gly Ile Val Ile Ile Ser Cys Leu Ser Asn Ile Ile Pro Ile Tyr Val  
305 310 315 320

Gln Thr Val Arg Gly Leu Gly Ala Ser Ile Ala Gly Leu Ile Leu Met  
325 330 335

Pro Ala Gly Ile Ile Lys Thr Ile Leu Ala Pro Ile Ser Gly Lys Leu  
340 345 350

Tyr Asp Lys Val Gly Val Ala Arg Ile Gly Leu Ile Gly Gly Ile Leu  
355 360 365

Leu Leu Val Gly Ser Leu Leu Leu Val Thr Leu Asn Glu Ala Ser Ser  
370 375 380

Leu Tyr Leu Leu Met Ile Tyr Tyr Gly Ile Leu Ser Ala Gly Phe Gly  
385 390 395 400

Leu Phe Asn Ile Pro Ile Thr Thr Ala Gly Met Asn Ile Met Ala Lys  
405 410 415

Glu Asp Met Gly His Ala Thr Ser Ala Arg Gln Thr Val Arg Gln Ile  
420 425 430

Ser Ser Ser Phe Ala Val Ser Leu Ser Phe Ile Ile Met Thr Leu Val  
435 440 445

Thr Ile Ala Thr Ser Gly Gln Ser Val Gly Val Phe Gln Asp Gly Gly  
450 455 460

Pro Thr Asp Leu Asn Met Ala Gly Val Arg Gly Ala Phe Ile Leu Val  
465 470 475 480

Ala Ile Phe Ser Ile Leu Ala Met Ile Leu Ile Phe Phe Leu Lys Asp  
485 490 495

Pro Lys Glu Lys Pro Asp Gln  
500

<210> 69  
 <211> 4392  
 <212> DNA  
 <213> *Alloioococcus otitidis*

<220>  
 <221> CDS  
 <222> (58)..(4392)  
 <223>

<400> 69  
 aagggttcag gactttcgta tctggccctt tttggctcat tagaaagcag ggcaaag 57

atg tca ctc aat caa aaa gaa atg tat caa gta ttg atg cag caa gtc 105  
 Met Ser Leu Asn Gln Lys Glu Met Tyr Gln Val Leu Met Gln Gln Val  
 1 5 10 15

cac tta gaa gaa cac cta caa gac cga ccc ctt ctt aaa gcc ggc agt 153  
 His Leu Glu Glu His Leu Gln Asp Arg Pro Leu Leu Lys Ala Gly Ser  
 20 25 30

ttg aag caa att gtt gtt tac aag gct caa caa gcc tgg gac ctg acc 201  
 Leu Lys Gln Ile Val Val Tyr Lys Ala Gln Gln Ala Trp Asp Leu Thr  
 35 40 45

ctc caa ttt cct cag atc ctc cct ttt aag gac ttc caa gtt ttg gag 249  
 Leu Gln Phe Pro Gln Ile Leu Pro Phe Lys Asp Phe Gln Val Leu Glu  
 50 55 60

tct gcc ctc ttg cag cat atc cca gaa gtc aac cag atc cat tta agg 297  
 Ser Ala Leu Leu Gln His Ile Pro Glu Val Asn Gln Ile His Leu Arg  
 65 70 75 80

gtt gat gcc caa gat gac agt ttt gac cag gac ctc ctc cag gac tat 345  
 Val Asp Ala Gln Asp Asp Ser Phe Asp Gln Asp Leu Leu Gln Asp Tyr  
 85 90 95

tgg cct aag gcg gtg aag ttt agc gga gtc gat tct ccc ctt tgc aat 393  
 Trp Pro Lys Ala Val Lys Phe Ser Gly Val Asp Ser Pro Leu Cys Asn  
 100 105 110

gac tta cta gac aag acc ctc cct tat cta gat ggg aag caa gtt tac 441  
 Asp Leu Leu Asp Lys Thr Leu Pro Tyr Leu Asp Gly Lys Gln Val Tyr  
 115 120 125

ttt gac ctg gac cat gaa gtg acc cgg gac aag ttt gac cat gac ttc 489  
 Phe Asp Leu Asp His Glu Val Thr Arg Asp Lys Phe Asp His Asp Phe  
 130 135 140

cta cct cgg atc caa gct ggc tac cag caa gtg ggc ttt ccc aac cac 537  
 Leu Pro Arg Ile Gln Ala Gly Tyr Gln Gln Val Gly Phe Pro Asn His  
 145 150 155 160

ttt aaa atc aag gct agg gtc gat gcc cag aaa aat tca gat caa att 585

Phe Lys Ile Lys Ala Arg Val Asp Ala Gln Lys Asn Ser Asp Gln Ile	
165 170 175	
gcc gcc ttc cgt aaa gaa aaa gaa gaa aaa gac cag gcc ttg tct caa	633
Ala Ala Phe Arg Lys Glu Lys Glu Glu Lys Asp Gln Ala Leu Ser Gln	
180 185 190	
gag cta acc aac caa ttt atc aag gcc agc caa aag aaa gaa gaa ggg	681
Glu Leu Thr Asn Gln Phe Ile Lys Ala Ser Gln Lys Lys Glu Glu Gly	
195 200 205	
gga tcc aaa gcc aag tcg gag gcc ttg aag atg ggc cgg gcc atc cct	729
Gly Ser Lys Ala Lys Ser Glu Ala Leu Lys Met Gly Arg Ala Ile Pro	
210 215 220	
gac cac gaa acg att acc cag atg gtt gat gtg gaa gaa gaa gag agc	777
Asp His Glu Thr Ile Thr Gln Met Val Asp Val Glu Glu Glu Glu Ser	
225 230 235 240	
cgt ctg acc ttt gaa gga tac gtt ttt gat gtg gaa atc aaa tcc ctc	825
Arg Leu Thr Phe Glu Gly Tyr Val Phe Asp Val Glu Ile Lys Ser Leu	
245 250 255	
cgg tca gat aga aag ctc ctt ctc ttt aaa atg acc gac tat agc tct	873
Arg Ser Asp Arg Lys Leu Leu Leu Phe Lys Met Thr Asp Tyr Ser Ser	
260 265 270	
tcc ttc cta ttc aaa aaa ttc tct aat aat tct tct gac gaa gcc cta	921
Ser Phe Leu Phe Lys Lys Phe Ser Asn Asn Ser Ser Asp Glu Ala Leu	
275 280 285	
ttt gac caa gtc caa gag gga atg tgg ctc aag gtt aga ggc agt gtt	969
Phe Asp Gln Val Gln Glu Gly Met Trp Leu Lys Val Arg Gly Ser Val	
290 295 300	
caa gaa gat acc ttt gtc aaa gac cta gtt gtc atg gcc caa gac atc	1017
Gln Glu Asp Thr Phe Val Lys Asp Leu Val Val Met Ala Gln Asp Ile	
305 310 315 320	
caa gag gtc aaa aaa gaa ccc cgg cgg gac ctg gct aag gaa ggg gag	1065
Gln Glu Val Lys Lys Glu Pro Arg Arg Asp Leu Ala Lys Glu Gly Glu	
325 330 335	
aag agg gtg gaa ctt cat gcc cat acc acc atg agt cag atg gac ggt	1113
Lys Arg Val Glu Leu His Ala His Thr Thr Met Ser Gln Met Asp Gly	
340 345 350	
ttg gtg ccg gcc aag gat ttg gtc aag caa gca gcc gct ttt gac caa	1161
Leu Val Pro Ala Lys Asp Leu Val Lys Gln Ala Ala Ala Phe Asp Gln	
355 360 365	
ccg gct att gcc atc act gat cat gct gta gtc caa tcc ttc cca gag	1209
Pro Ala Ile Ala Ile Thr Asp His Ala Val Val Gln Ser Phe Pro Glu	
370 375 380	
gcc cat tat gct ggc tta gac act ggt gtt aaa att ctt tac ggt gtg	1257
Ala His Tyr Ala Gly Leu Asp Thr Gly Val Lys Ile Leu Tyr Gly Val	

385	390	395	400	
gaa gcc aat ttg gtt agt gat ggc gaa ttg gta gca tac aat ccg gcc				1305
Glu Ala Asn Leu Val Ser Asp Gly Glu Leu Val Ala Tyr Asn Pro Ala	405	410	415	
gat ata aag ctg gaa gag gca act tat gtg gtc ttc gac gtg gaa aca				1353
Asp Ile Lys Leu Glu Glu Ala Thr Tyr Val Val Phe Asp Val Glu Thr	420	425	430	
acc gga cta tcg gct cgt tat gac caa atc att gaa ttg gcc gct gtg				1401
Thr Gly Leu Ser Ala Arg Tyr Asp Gln Ile Ile Glu Leu Ala Ala Val	435	440	445	
aag atg gaa aat ggg gaa atc gtt tct gaa ttc caa gaa ttt att gac				1449
Lys Met Glu Asn Gly Glu Ile Val Ser Glu Phe Gln Glu Phe Ile Asp	450	455	460	
cca ggc cag ccc ttg tct gag act acg acc aat ttg acc ggg atc acc				1497
Pro Gly Gln Pro Leu Ser Glu Thr Thr Thr Asn Leu Thr Gly Ile Thr	465	470	475	480
gat gac atg gtc caa gga tcc aaa agt gaa gac gaa gtc ctc cat gcc				1545
Asp Asp Met Val Gln Gly Ser Lys Ser Glu Asp Glu Val Leu His Ala	485	490	495	
ttt caa gcc ttt tca gaa ggc act gtc ttg gtc gcc cat aac gct tcc				1593
Phe Gln Ala Phe Ser Glu Gly Thr Val Leu Val Ala His Asn Ala Ser	500	505	510	
ttt gac atg ggc ttt atc aat acg gcc tac caa cga cat ggc cta gga				1641
Phe Asp Met Gly Phe Ile Asn Thr Ala Tyr Gln Arg His Gly Leu Gly	515	520	525	
caa gct gac cag cct gtg att gat acc ttg gaa ttg tcc cgc atg ctc				1689
Gln Ala Asp Gln Pro Val Ile Asp Thr Leu Glu Leu Ser Arg Met Leu	530	535	540	
cac cca aac ttg aaa agc cac cgg tta aac act ctg gct aag cgg tat				1737
His Pro Asn Leu Lys Ser His Arg Leu Asn Thr Leu Ala Lys Arg Tyr	545	550	555	560
gac gtg gcc tta gaa cac cac cac cgg gcc atc tat gac tcg gag tca				1785
Asp Val Ala Leu Glu His His His Arg Ala Ile Tyr Asp Ser Glu Ser	565	570	575	
acg gct aaa ctc ttg tgg atc ttc tta aaa gaa gcc aaa gac caa tat				1833
Thr Ala Lys Leu Leu Trp Ile Phe Leu Lys Glu Ala Lys Asp Gln Tyr	580	585	590	
gac atg act agc cac caa gac ttg aat agc cag gtg ggg gaa ggc gag				1881
Asp Met Thr Ser His Gln Asp Leu Asn Ser Gln Val Gly Glu Gly Glu	595	600	605	
gct tac aag cag gcc cgg cca acc cat gcc agt att ttg gtc aag aat				1929
Ala Tyr Lys Gln Ala Arg Pro Thr His Ala Ser Ile Leu Val Lys Asn	610	615	620	

caa aaa ggc ttg aaa aac ctc ttt aaa att gtc tcc cac gcc cat gtc Gln Lys Gly Leu Lys Asn Leu Phe Lys Ile Val Ser His Ala His Val 625 630 635 640	1977
aac tac ttc tac cgg gtt ccc cgt ata cct aag tct atc ttg agc aag Asn Tyr Phe Tyr Arg Val Pro Arg Ile Pro Lys Ser Ile Leu Ser Lys 645 650 655	2025
tac cgg gaa ggc ctt ttg gtt ggg tct ggt tgc gga cag gga gag ctc Tyr Arg Glu Gly Leu Leu Val Gly Ser Gly Cys Gly Gln Gly Glu Leu 660 665 670	2073
ttt gag gct att atg caa aag ggc tat gac gaa gcc ttg gca gtt gcc Phe Glu Ala Ile Met Gln Lys Gly Tyr Asp Glu Ala Leu Ala Val Ala 675 680 685	2121
cag gac tat gat tat att gaa gtt atg ccc aag tca gcc tat att gac Gln Asp Tyr Asp Tyr Ile Glu Val Met Pro Lys Ser Ala Tyr Ile Asp 690 695 700	2169
ctc ttg gac cgg gac tta atc aag gat gag gca acc ctt gaa gaa atg Leu Leu Asp Arg Asp Leu Ile Lys Asp Glu Ala Thr Leu Glu Glu Met 705 710 715 720	2217
att gaa aac ctg gtt aaa ata ggc cat gaa ctt gat ata ccc gtg gta Ile Glu Asn Leu Val Lys Ile Gly His Glu Leu Asp Ile Pro Val Val 725 730 735	2265
gct aca ggg aat gtc cac tac cta aac cca gaa gat gcc gtt tta cgg Ala Thr Gly Asn Val His Tyr Leu Asn Pro Glu Asp Ala Val Leu Arg 740 745 750	2313
gat atc ctc ctg gaa act gcc aaa aag gga gcc ttc tcc aaa gcc cgg Asp Ile Leu Leu Glu Thr Ala Lys Lys Gly Ala Phe Ser Lys Ala Arg 755 760 765	2361
aac cca gaa gtc cac ttt aga aca aca gat gaa atg tta gaa gag ttt Asn Pro Glu Val His Phe Arg Thr Thr Asp Glu Met Leu Glu Glu Phe 770 775 780	2409
tcc ttc cta ggc cag gac cag gct tat gag att gtg gtc acc aac acc Ser Phe Leu Gly Gln Asp Gln Ala Tyr Glu Ile Val Val Thr Asn Thr 785 790 795 800	2457
caa aaa att gct gat tct atc gaa tca atc tct cct gtc aag gaa ggc Gln Lys Ile Ala Asp Ser Ile Glu Ser Ile Ser Pro Val Lys Glu Gly 805 810 815	2505
ctc tat gcc ccg aaa atg gaa ggg tcg gac caa gag ata cgt cag atg Leu Tyr Ala Pro Lys Met Glu Gly Ser Asp Gln Glu Ile Arg Gln Met 820 825 830	2553
agt tac aag caa gcc aag gct ctc tat ggc gac ccc ttg cca agt att Ser Tyr Lys Gln Ala Lys Ala Leu Tyr Gly Asp Pro Leu Pro Ser Ile 835 840 845	2601

gta gag gaa agg ctc gaa aaa gag ttg aag agt att att gac aac aat Val Glu Glu Arg Leu Glu Lys Glu Leu Lys Ser Ile Ile Asp Asn Asn 850 855 860	2649
ttc tct gtc att tac tta att tcc cag aaa ttg gtc aaa aaa agt gtt Phe Ser Val Ile Tyr Leu Ile Ser Gln Lys Leu Val Lys Lys Ser Val 865 870 875 880	2697
gaa gat ggc tat ttg gtt ggt tcc agg ggg tcg gtt ggg tca agc ttt Glu Asp Gly Tyr Leu Val Gly Ser Arg Gly Ser Val Gly Ser Ser Phe 885 890 895	2745
gtg gcc acc atg acc ggg atc aca gaa gtc aac cca cta ccg ccc cac Val Ala Thr Met Thr Gly Ile Thr Glu Val Asn Pro Leu Pro Pro His 900 905 910	2793
tac cgc tgt cct aac tgc cag cac acc gaa ttc ttc aca aat ggg gaa Tyr Arg Cys Pro Asn Cys Gln His Thr Glu Phe Phe Thr Asn Gly Glu 915 920 925	2841
gtg ggg tcc ggc ttt gac tta gag gcc aaa aaa tgt ccg gaa tgt caa Val Gly Ser Gly Phe Asp Leu Glu Ala Lys Lys Cys Pro Glu Cys Gln 930 935 940	2889
agc cta atg gaa tca gac ggc cac gac att ccc ttc gaa acc ttc ctt Ser Leu Met Glu Ser Asp Gly His Asp Ile Pro Phe Glu Thr Phe Leu 945 950 955 960	2937
ggg ttt aat ggg gac aag gtg cca gat atc gat ttg aac ttc tca ggt Gly Phe Asn Gly Asp Lys Val Pro Asp Ile Asp Leu Asn Phe Ser Gly 965 970 975	2985
gaa tac cag gcc aag gcc cac aac tat acc aag gtt ttg ttt gga gaa Glu Tyr Gln Ala Lys Ala His Asn Tyr Thr Lys Val Leu Phe Gly Glu 980 985 990	3033
gac cat gtc tac cgg gca ggg acc atc acg acg att gct gac aag acg Asp His Val Tyr Arg Ala Gly Thr Ile Thr Thr Ile Ala Asp Lys Thr 995 1000 1005	3081
gcc ttt ggt ttt gtc aag ggt tat gaa agg gac aag cag ata aac tac Ala Phe Gly Phe Val Lys Gly Tyr Glu Arg Asp Lys Gln Ile Asn Tyr 1010 1015 1020	3129
cgg tcg gct gaa gtg gac cgg ctg tca gat ggt tta acc gga gtg aga Arg Ser Ala Glu Val Asp Arg Leu Ser Asp Gly Leu Thr Gly Val Arg 1025 1030 1035 1040	3177
cgg tca acc ggc cag cac cca gga ggg att atc gtc ata ccg gat gac Arg Ser Thr Gly Gln His Pro Gly Gly Ile Ile Val Ile Pro Asp Asp 1045 1050 1055	3225
atg gat gtg ttt gat ttc acc ccc atc cag tac ccg gct gac gac cag Met Asp Val Phe Asp Phe Thr Pro Ile Gln Tyr Pro Ala Asp Asp Gln 1060 1065 1070	3273
acg gct gag tgg caa act acc cac ttt gac ttc cac tcc atc gac gaa	3321

Thr Ala Glu Trp Gln Thr Thr His Phe Asp Phe His Ser Ile Asp Glu 1075 1080 1085	
aac gtc ttg aag ctg gat atc ctg gga cat gat gac ccg acc atg atc Asn Val Leu Lys Leu Asp Ile Leu Gly His Asp Asp Pro Thr Met Ile 1090 1095 1100	3369
cga aaa ctc cag gac ttg tcc ggc ttt gac cct caa gaa ata ccg gta Arg Lys Leu Gln Asp Leu Ser Gly Phe Asp Pro Gln Glu Ile Pro Val 1105 1110 1115 1120	3417
agt gat gaa gat gtt atg aaa att ttc tca ggc ccg gaa gtt cta ggg Ser Asp Glu Asp Val Met Lys Ile Phe Ser Gly Pro Glu Val Leu Gly 1125 1130 1135	3465
gtg acc cca gag caa att ttc tcc aat acc gga act ctc gga gta cct Val Thr Pro Glu Gln Ile Phe Ser Asn Thr Gly Thr Leu Gly Val Pro 1140 1145 1150	3513
gaa ttt ggt acc caa ttt gtc cga gaa atg tta gag caa acc cac ccc Glu Phe Gly Thr Gln Phe Val Arg Glu Met Leu Glu Gln Thr His Pro 1155 1160 1165	3561
tct acc ttt gct gaa ctc ttg cag atc tca ggc ctc tcc cac ggg aca Ser Thr Phe Ala Glu Leu Leu Gln Ile Ser Gly Leu Ser His Gly Thr 1170 1175 1180	3609
gat gtt tgg ctg ggc aat gct gaa gaa tta att cgc aac cac aac att Asp Val Trp Leu Gly Asn Ala Glu Glu Leu Ile Arg Asn His Asn Ile 1185 1190 1195 1200	3657
ccc ttg tcc gag gtg atc ggc tgc cgg gat gat atc atg gtc tac ctt Pro Leu Ser Glu Val Ile Gly Cys Arg Asp Asp Ile Met Val Tyr Leu 1205 1210 1215	3705
caa cac caa ggt ctt gaa gac agc ctg gcc ttt aag att atg gaa ttt Gln His Gln Gly Leu Glu Asp Ser Leu Ala Phe Lys Ile Met Glu Phe 1220 1225 1230	3753
gtt cgt aag ggt cgg ggc ttg caa gat gac tgg att gct acc atg aaa Val Arg Lys Gly Arg Gly Leu Gln Asp Asp Trp Ile Ala Thr Met Lys 1235 1240 1245	3801
gaa aat gat gtt cct gat tgg tat att gaa tcc tgc aaa aaa atc aag Glu Asn Asp Val Pro Asp Trp Tyr Ile Glu Ser Cys Lys Lys Ile Lys 1250 1255 1260	3849
tac atg ttc cct aaa gcc cac gca gct gcc tat gtc ttg atg gcc ctt Tyr Met Phe Pro Lys Ala His Ala Ala Tyr Val Leu Met Ala Leu 1265 1270 1275 1280	3897
agg gta gct tac ttt aaa gtc cac tac ccc ctt tac tac tac gct gcc Arg Val Ala Tyr Phe Lys Val His Tyr Pro Leu Tyr Tyr Tyr Ala Ala 1285 1290 1295	3945
tac ttt tcc atc cgg gct agt gat ttt gac tta att gct atg gtc aag Tyr Phe Ser Ile Arg Ala Ser Asp Phe Asp Leu Ile Ala Met Val Lys	3993



1300	1305	1310	
ggc aag gaa ggc att aaa ggg gct atg aag gaa atc agg gac aag gaa Gly Lys Glu Gly Ile Lys Gly Ala Met Lys Glu Ile Arg Asp Lys Glu 1315 1320 1325			4041
aga gaa aaa act gcc aca gct aag gac aaa gcc ttg ctc acc gtc ctt Arg Glu Lys Thr Ala Thr Ala Lys Asp Lys Ala Leu Leu Thr Val Leu 1330 1335 1340			4089
gaa gta gcc aat gaa atg gtt gaa cgg ggt ttt gac ttc aag atg gtg Glu Val Ala Asn Glu Met Val Glu Arg Gly Phe Asp Phe Lys Met Val 1345 1350 1355 1360			4137
gac atc aac aag tcc caa gcc aaa gac ttt gtc atc gaa gac aat ggc Asp Ile Asn Lys Ser Gln Ala Lys Asp Phe Val Ile Glu Asp Asn Gly 1365 1370 1375			4185
ctt cgt gct cca ttt agg gca gtc cct tcc ttg ggg tcc agt gcc gcc Leu Arg Ala Pro Phe Arg Ala Val Pro Ser Leu Gly Ser Ser Ala Ala 1380 1385 1390			4233
cag gct gtc att gat gcc agg gag gac agc gac ttc ttg tcc aag gaa Gln Ala Val Ile Asp Ala Arg Glu Asp Ser Asp Phe Leu Ser Lys Glu 1395 1400 1405			4281
gac cta tca aaa cgg ggc aag ttg tcg aaa acg gtc atg gac tac ctg Asp Leu Ser Lys Arg Gly Lys Leu Ser Lys Thr Val Met Asp Tyr Leu 1410 1415 1420			4329
gac aat aac cac gtt tta gac cac ctg ccg gac gaa aac caa ctt tcc Asp Asn Asn His Val Leu Asp His Leu Pro Asp Glu Asn Gln Leu Ser 1425 1430 1435 1440			4377
ctc ttt gac ttt taa Leu Phe Asp Phe			4392

&lt;210&gt; 70

&lt;211&gt; 1444

&lt;212&gt; PRT

&lt;213&gt; Alloiococcus otitidis

&lt;400&gt; 70

Met Ser Leu Asn Gln Lys Glu Met Tyr Gln Val Leu Met Gln Gln Val
1 5 10 15

His Leu Glu Glu His Leu Gln Asp Arg Pro Leu Leu Lys Ala Gly Ser
20 25 30

Leu Lys Gln Ile Val Val Tyr Lys Ala Gln Gln Ala Trp Asp Leu Thr
35 40 45

Leu Gln Phe Pro Gln Ile Leu Pro Phe Lys Asp Phe Gln Val Leu Glu

50

55

60

Ser Ala Leu Leu Gln His Ile Pro Glu Val Asn Gln Ile His Leu Arg  
65 70 75 80

Val Asp Ala Gln Asp Asp Ser Phe Asp Gln Asp Leu Leu Gln Asp Tyr  
85 90 95

Trp Pro Lys Ala Val Lys Phe Ser Gly Val Asp Ser Pro Leu Cys Asn  
100 105 110

Asp Leu Leu Asp Lys Thr Leu Pro Tyr Leu Asp Gly Lys Gln Val Tyr  
115 120 125

Phe Asp Leu Asp His Glu Val Thr Arg Asp Lys Phe Asp His Asp Phe  
130 135 140

Leu Pro Arg Ile Gln Ala Gly Tyr Gln Gln Val Gly Phe Pro Asn His  
145 150 155 160

Phe Lys Ile Lys Ala Arg Val Asp Ala Gln Lys Asn Ser Asp Gln Ile  
165 170 175

Ala Ala Phe Arg Lys Glu Lys Glu Glu Lys Asp Gln Ala Leu Ser Gln  
180 185 190

Glu Leu Thr Asn Gln Phe Ile Lys Ala Ser Gln Lys Lys Glu Glu Gly  
195 200 205

Gly Ser Lys Ala Lys Ser Glu Ala Leu Lys Met Gly Arg Ala Ile Pro  
210 215 220

Asp His Glu Thr Ile Thr Gln Met Val Asp Val Glu Glu Glu Glu Ser  
225 230 235 240

Arg Leu Thr Phe Glu Gly Tyr Val Phe Asp Val Glu Ile Lys Ser Leu  
245 250 255

Arg Ser Asp Arg Lys Leu Leu Leu Phe Lys Met Thr Asp Tyr Ser Ser  
260 265 270

Ser Phe Leu Phe Lys Lys Phe Ser Asn Asn Ser Ser Asp Glu Ala Leu  
275 280 285

Phe Asp Gln Val Gln Glu Gly Met Trp Leu Lys Val Arg Gly Ser Val  
 290 295 300

Gln Glu Asp Thr Phe Val Lys Asp Leu Val Val Met Ala Gln Asp Ile  
 305 310 315 320

Gln Glu Val Lys Lys Glu Pro Arg Arg Asp Leu Ala Lys Glu Gly Glu  
 325 330 335

Lys Arg Val Glu Leu His Ala His Thr Thr Met Ser Gln Met Asp Gly  
 340 345 350

Leu Val Pro Ala Lys Asp Leu Val Lys Gln Ala Ala Ala Phe Asp Gln  
 355 360 365

Pro Ala Ile Ala Ile Thr Asp His Ala Val Val Gln Ser Phe Pro Glu  
 370 375 380

Ala His Tyr Ala Gly Leu Asp Thr Gly Val Lys Ile Leu Tyr Gly Val  
 385 390 395 400

Glu Ala Asn Leu Val Ser Asp Gly Glu Leu Val Ala Tyr Asn Pro Ala  
 405 410 415

Asp Ile Lys Leu Glu Glu Ala Thr Tyr Val Val Phe Asp Val Glu Thr  
 420 425 430

Thr Gly Leu Ser Ala Arg Tyr Asp Gln Ile Ile Glu Leu Ala Ala Val  
 435 440 445

Lys Met Glu Asn Gly Glu Ile Val Ser Glu Phe Gln Glu Phe Ile Asp  
 450 455 460

Pro Gly Gln Pro Leu Ser Glu Thr Thr Thr Asn Leu Thr Gly Ile Thr  
 465 470 475 480

Asp Asp Met Val Gln Gly Ser Lys Ser Glu Asp Glu Val Leu His Ala  
 485 490 495

Phe Gln Ala Phe Ser Glu Gly Thr Val Leu Val Ala His Asn Ala Ser  
 500 505 510

Phe Asp Met Gly Phe Ile Asn Thr Ala Tyr Gln Arg His Gly Leu Gly  
515 520 525

Gln Ala Asp Gln Pro Val Ile Asp Thr Leu Glu Leu Ser Arg Met Leu  
530 535 540

His Pro Asn Leu Lys Ser His Arg Leu Asn Thr Leu Ala Lys Arg Tyr  
545 550 555 560

Asp Val Ala Leu Glu His His His Arg Ala Ile Tyr Asp Ser Glu Ser  
565 570 575

Thr Ala Lys Leu Leu Trp Ile Phe Leu Lys Glu Ala Lys Asp Gln Tyr  
580 585 590

Asp Met Thr Ser His Gln Asp Leu Asn Ser Gln Val Gly Glu Gly Glu  
595 600 605

Ala Tyr Lys Gln Ala Arg Pro Thr His Ala Ser Ile Leu Val Lys Asn  
610 615 620

Gln Lys Gly Leu Lys Asn Leu Phe Lys Ile Val Ser His Ala His Val  
625 630 635 640

Asn Tyr Phe Tyr Arg Val Pro Arg Ile Pro Lys Ser Ile Leu Ser Lys  
645 650 655

Tyr Arg Glu Gly Leu Leu Val Gly Ser Gly Cys Gly Gln Gly Glu Leu  
660 665 670

Phe Glu Ala Ile Met Gln Lys Gly Tyr Asp Glu Ala Leu Ala Val Ala  
675 680 685

Gln Asp Tyr Asp Tyr Ile Glu Val Met Pro Lys Ser Ala Tyr Ile Asp  
690 695 700

Leu Leu Asp Arg Asp Leu Ile Lys Asp Glu Ala Thr Leu Glu Glu Met  
705 710 715 720

Ile Glu Asn Leu Val Lys Ile Gly His Glu Leu Asp Ile Pro Val Val  
725 730 735

Ala Thr Gly Asn Val His Tyr Leu Asn Pro Glu Asp Ala Val Leu Arg  
740 745 750

Asp Ile Leu Leu Glu Thr Ala Lys Lys Gly Ala Phe Ser Lys Ala Arg  
755 760 765

Asn Pro Glu Val His Phe Arg Thr Thr Asp Glu Met Leu Glu Glu Phe  
770 775 780

Ser Phe Leu Gly Gln Asp Gln Ala Tyr Glu Ile Val Val Thr Asn Thr  
785 790 795 800

Gln Lys Ile Ala Asp Ser Ile Glu Ser Ile Ser Pro Val Lys Glu Gly  
805 810 815

Leu Tyr Ala Pro Lys Met Glu Gly Ser Asp Gln Glu Ile Arg Gln Met  
820 825 830

Ser Tyr Lys Gln Ala Lys Ala Leu Tyr Gly Asp Pro Leu Pro Ser Ile  
835 840 845

Val Glu Glu Arg Leu Glu Lys Glu Leu Lys Ser Ile Ile Asp Asn Asn  
850 855 860

Phe Ser Val Ile Tyr Leu Ile Ser Gln Lys Leu Val Lys Lys Ser Val  
865 870 875 880

Glu Asp Gly Tyr Leu Val Gly Ser Arg Gly Ser Val Gly Ser Ser Phe  
885 890 895

Val Ala Thr Met Thr Gly Ile Thr Glu Val Asn Pro Leu Pro Pro His  
900 905 910

Tyr Arg Cys Pro Asn Cys Gln His Thr Glu Phe Phe Thr Asn Gly Glu  
915 920 925

Val Gly Ser Gly Phe Asp Leu Glu Ala Lys Lys Cys Pro Glu Cys Gln  
930 935 940

Ser Leu Met Glu Ser Asp Gly His Asp Ile Pro Phe Glu Thr Phe Leu  
945 950 955 960

Gly Phe Asn Gly Asp Lys Val Pro Asp Ile Asp Leu Asn Phe Ser Gly

965

970

975

Glu Tyr Gln Ala Lys Ala His Asn Tyr Thr Lys Val Leu Phe Gly Glu  
980 985 990

Asp His Val Tyr Arg Ala Gly Thr Ile Thr Thr Ile Ala Asp Lys Thr  
995 1000 1005

Ala Phe Gly Phe Val Lys Gly Tyr Glu Arg Asp Lys Gln Ile Asn Tyr  
1010 1015 1020

Arg Ser Ala Glu Val Asp Arg Leu Ser Asp Gly Leu Thr Gly Val Arg  
1025 1030 1035 1040

Arg Ser Thr Gly Gln His Pro Gly Gly Ile Ile Val Ile Pro Asp Asp  
1045 1050 1055

Met Asp Val Phe Asp Phe Thr Pro Ile Gln Tyr Pro Ala Asp Asp Gln  
1060 1065 1070

Thr Ala Glu Trp Gln Thr Thr His Phe Asp Phe His Ser Ile Asp Glu  
1075 1080 1085

Asn Val Leu Lys Leu Asp Ile Leu Gly His Asp Asp Pro Thr Met Ile  
1090 1095 1100

Arg Lys Leu Gln Asp Leu Ser Gly Phe Asp Pro Gln Glu Ile Pro Val  
1105 1110 1115 1120

Ser Asp Glu Asp Val Met Lys Ile Phe Ser Gly Pro Glu Val Leu Gly  
1125 1130 1135

Val Thr Pro Glu Gln Ile Phe Ser Asn Thr Gly Thr Leu Gly Val Pro  
1140 1145 1150

Glu Phe Gly Thr Gln Phe Val Arg Glu Met Leu Glu Gln Thr His Pro  
1155 1160 1165

Ser Thr Phe Ala Glu Leu Leu Gln Ile Ser Gly Leu Ser His Gly Thr  
1170 1175 1180

Asp Val Trp Leu Gly Asn Ala Glu Glu Leu Ile Arg Asn His Asn Ile  
1185 1190 1195 1200

Pro Leu Ser Glu Val Ile Gly Cys Arg Asp Asp Ile Met Val Tyr Leu  
1205 1210 1215

Gln His Gln Gly Leu Glu Asp Ser Leu Ala Phe Lys Ile Met Glu Phe  
1220 1225 1230

Val Arg Lys Gly Arg Gly Leu Gln Asp Asp Trp Ile Ala Thr Met Lys  
1235 1240 1245

Glu Asn Asp Val Pro Asp Trp Tyr Ile Glu Ser Cys Lys Lys Ile Lys  
1250 1255 1260

Tyr Met Phe Pro Lys Ala His Ala Ala Ala Tyr Val Leu Met Ala Leu  
1265 1270 1275 1280

Arg Val Ala Tyr Phe Lys Val His Tyr Pro Leu Tyr Tyr Tyr Ala Ala  
1285 1290 1295

Tyr Phe Ser Ile Arg Ala Ser Asp Phe Asp Leu Ile Ala Met Val Lys  
1300 1305 1310

Gly Lys Glu Gly Ile Lys Gly Ala Met Lys Glu Ile Arg Asp Lys Glu  
1315 1320 1325

Arg Glu Lys Thr Ala Thr Ala Lys Asp Lys Ala Leu Leu Thr Val Leu  
1330 1335 1340

Glu Val Ala Asn Glu Met Val Glu Arg Gly Phe Asp Phe Lys Met Val  
1345 1350 1355 1360

Asp Ile Asn Lys Ser Gln Ala Lys Asp Phe Val Ile Glu Asp Asn Gly  
1365 1370 1375

Leu Arg Ala Pro Phe Arg Ala Val Pro Ser Leu Gly Ser Ser Ala Ala  
1380 1385 1390

Gln Ala Val Ile Asp Ala Arg Glu Asp Ser Asp Phe Leu Ser Lys Glu  
1395 1400 1405

Asp Leu Ser Lys Arg Gly Lys Leu Ser Lys Thr Val Met Asp Tyr Leu  
1410 1415 1420

Asp Asn Asn His Val Leu Asp His Leu Pro Asp Glu Asn Gln Leu Ser  
 1425 1430 1435 1440

Leu Phe Asp Phe

<210> 71  
 <211> 1326  
 <212> DNA  
 <213> *Alloioococcus otitidis*

<220>  
 <221> CDS  
 <222> (19)..(1326)  
 <223>

<400> 71  
 aagaaaggag gactcaat atg tct atg ttt gtc gac tac acc aaa gtt aac 51  
                           Met Ser Met Phe Val Asp Tyr Thr Lys Val Asn  
                           1                          5                          10

ctg aga gcc ggt aag ggc ggt gac gga atg gtg gct ttt aga cga gaa 99  
 Leu Arg Ala Gly Lys Gly Gly Asp Gly Met Val Ala Phe Arg Arg Glu  
                           15                          20                          25

aag tat gag ccc aat ggt gga cca gca ggc ggc gac ggt ggc agt ggc 147  
 Lys Tyr Glu Pro Asn Gly Gly Pro Ala Gly Gly Asp Gly Gly Ser Gly  
                           30                          35                          40

ggt aac att atc ttc aag gta gat gaa ggc ctc cgt acc ctg gta gac 195  
 Gly Asn Ile Ile Phe Lys Val Asp Glu Gly Leu Arg Thr Leu Val Asp  
                           45                          50                          55

ttc cgc tac aac ccc cat ttt aag gca gat agt ggc caa aat ggt atg 243  
 Phe Arg Tyr Asn Pro His Phe Lys Ala Asp Ser Gly Gln Asn Gly Met  
                           60                          65                          70                          75

ccc aag ggg atg aat ggt aag aag gca gag gac ttg att atc agt gtc 291  
 Pro Lys Gly Met Asn Gly Lys Lys Ala Glu Asp Leu Ile Ile Ser Val  
                           80                          85                          90

ccg cct gga acc att atc cgg gat gcc caa agt aag gct ata ctt gct 339  
 Pro Pro Gly Thr Ile Ile Arg Asp Ala Gln Ser Lys Ala Ile Leu Ala  
                           95                          100                          105

gac tta caa gaa gaa gga caa gaa gtc ttg gca gcc caa ggt ggc cgg 387  
 Asp Leu Gln Glu Glu Gly Gln Val Leu Ala Ala Gln Gly Gly Arg  
                           110                          115                          120

gga ggt cgg ggc aat aaa cgt ttt gct acg cat aag aac cca gca ccc 435  
 Gly Gly Arg Gly Asn Lys Arg Phe Ala Thr His Lys Asn Pro Ala Pro  
                           125                          130                          135

tcc att gcc gaa aac ggc gag ccg ggc caa gag cgg gat gtc gaa ttg 483  
 Ser Ile Ala Glu Asn Gly Glu Pro Gly Gln Glu Arg Asp Val Glu Leu



140	145	150	155	
gaa tta aaa gtc atg gcc gat gtt ggc cta gtg ggt tat cct tct gtc				531
Glu Leu Lys Val Met Ala Asp Val Gly Leu Val Gly Tyr Pro Ser Val				
	160	165	170	
ggg aaa tcg acc ctt ttg tcg gtt gtc tca ggc gct aaa ccc aaa att				579
Gly Lys Ser Thr Leu Leu Ser Val Val Ser Gly Ala Lys Pro Lys Ile				
	175	180	185	
gga gcc tat cac ttt act aca ctt gcc cct aat tta ggt gta gtg aat				627
Gly Ala Tyr His Phe Thr Thr Leu Ala Pro Asn Leu Gly Val Val Asn				
	190	195	200	
gca gtg gac ggc aag gaa ttt gtc ttg gcg gat att cct ggc tta att				675
Ala Val Asp Gly Lys Glu Phe Val Leu Ala Asp Ile Pro Gly Leu Ile				
	205	210	215	
gaa ggg gct tca gaa ggg gtt ggt ttg ggg att gac ttc ctc aag cat				723
Glu Gly Ala Ser Glu Gly Val Gly Leu Gly Ile Asp Phe Leu Lys His				
	220	225	230	235
att gaa aga acc cgc atc ctc ctt cat gta ctt gat atg agc gga atg				771
Ile Glu Arg Thr Arg Ile Leu Leu His Val Leu Asp Met Ser Gly Met				
	240	245	250	
gaa ggt cgc cat cca att gat gat ttt gac cag att aac caa gaa cta				819
Glu Gly Arg His Pro Ile Asp Asp Phe Asp Gln Ile Asn Gln Glu Leu				
	255	260	265	
aaa gac tat aat gag aaa tta ttg gac cgc aag cag gtc att gtg gcc				867
Lys Asp Tyr Asn Glu Lys Leu Leu Asp Arg Lys Gln Val Ile Val Ala				
	270	275	280	
aat aaa atg gac ctg ccc cag tcc cgg gat aat tta atc gaa ttt aaa				915
Asn Lys Met Asp Leu Pro Gln Ser Arg Asp Asn Leu Ile Glu Phe Lys				
	285	290	295	
gcc gag tta gac agc cgg gac ctt gac tat gaa atc ttt gaa gtg tca				963
Ala Glu Leu Asp Ser Arg Asp Leu Asp Tyr Glu Ile Phe Glu Val Ser				
	300	305	310	315
gct gcc acc cag gct ggc att cag gac cta gtc atc cga cta gcc gac				1011
Ala Ala Thr Gln Ala Gly Ile Gln Asp Leu Val Ile Arg Leu Ala Asp				
	320	325	330	
tta gtc gac caa ctg gac caa gcc cca agt tta gac cag gaa gaa act				1059
Leu Val Asp Gln Leu Asp Gln Ala Pro Ser Leu Asp Gln Glu Glu Thr				
	335	340	345	
agt gaa gcc gac caa aga gtg gtc tac aag ttt caa gct gac caa gac				1107
Ser Glu Ala Asp Gln Arg Val Val Tyr Lys Phe Gln Ala Asp Gln Asp				
	350	355	360	
aaa ttt gac ctt gac cgc gac cct gaa ggg gta tgg ttg gtt tct ggt				1155
Lys Phe Asp Leu Asp Arg Asp Pro Glu Gly Val Trp Leu Val Ser Gly				
	365	370	375	

ccc aag gtt gag cgt ttg tat gcc atg acc aat ttt gac cac gag gaa 1203  
 Pro Lys Val Glu Arg Leu Tyr Ala Met Thr Asn Phe Asp His Glu Glu  
 380 385 390 395

gcc att atg cgg ttt tct cgc cag cta aga ggg atg gga gta gac caa 1251  
 Ala Ile Met Arg Phe Ser Arg Gln Leu Arg Gly Met Gly Val Asp Gln  
 400 405 410

gcc tta aga gac aag ggg gct cag tct ggt gac ctc gtc caa gtt gaa 1299  
 Ala Leu Arg Asp Lys Gly Ala Gln Ser Gly Asp Leu Val Gln Val Glu  
 415 420 425

gat ttt gtc ttt gag ttc atg gat tag 1326  
 Asp Phe Val Phe Glu Phe Met Asp  
 430 435

&lt;210&gt; 72

&lt;211&gt; 435

&lt;212&gt; PRT

<213> *Alloiococcus otitidis*

&lt;400&gt; 72

Met Ser Met Phe Val Asp Tyr Thr Lys Val Asn Leu Arg Ala Gly Lys  
 1 5 10 15

Gly Gly Asp Gly Met Val Ala Phe Arg Arg Glu Lys Tyr Glu Pro Asn  
 20 25 30

Gly Gly Pro Ala Gly Gly Asp Gly Gly Ser Gly Gly Asn Ile Ile Phe  
 35 40 45

Lys Val Asp Glu Gly Leu Arg Thr Leu Val Asp Phe Arg Tyr Asn Pro  
 50 55 60

His Phe Lys Ala Asp Ser Gly Gln Asn Gly Met Pro Lys Gly Met Asn  
 65 70 75 80

Gly Lys Lys Ala Glu Asp Leu Ile Ile Ser Val Pro Pro Gly Thr Ile  
 85 90 95

Ile Arg Asp Ala Gln Ser Lys Ala Ile Leu Ala Asp Leu Gln Glu Glu  
 100 105 110

Gly Gln Glu Val Leu Ala Ala Gln Gly Gly Arg Gly Gly Arg Gly Asn  
 115 120 125

Lys Arg Phe Ala Thr His Lys Asn Pro Ala Pro Ser Ile Ala Glu Asn

130

135

140

Gly Glu Pro Gly Gln Glu Arg Asp Val Glu Leu Glu Leu Lys Val Met  
 145 150 155 160

Ala Asp Val Gly Leu Val Gly Tyr Pro Ser Val Gly Lys Ser Thr Leu  
 165 170 175

Leu Ser Val Val Ser Gly Ala Lys Pro Lys Ile Gly Ala Tyr His Phe  
 180 185 190

Thr Thr Leu Ala Pro Asn Leu Gly Val Val Asn Ala Val Asp Gly Lys  
 195 200 205

Glu Phe Val Leu Ala Asp Ile Pro Gly Leu Ile Glu Gly Ala Ser Glu  
 210 215 220

Gly Val Gly Leu Gly Ile Asp Phe Leu Lys His Ile Glu Arg Thr Arg  
 225 230 235 240

Ile Leu Leu His Val Leu Asp Met Ser Gly Met Glu Gly Arg His Pro  
 245 250 255

Ile Asp Asp Phe Asp Gln Ile Asn Gln Glu Leu Lys Asp Tyr Asn Glu  
 260 265 270

Lys Leu Leu Asp Arg Lys Gln Val Ile Val Ala Asn Lys Met Asp Leu  
 275 280 285

Pro Gln Ser Arg Asp Asn Leu Ile Glu Phe Lys Ala Glu Leu Asp Ser  
 290 295 300

Arg Asp Leu Asp Tyr Glu Ile Phe Glu Val Ser Ala Ala Thr Gln Ala  
 305 310 315 320

Gly Ile Gln Asp Leu Val Ile Arg Leu Ala Asp Leu Val Asp Gln Leu  
 325 330 335

Asp Gln Ala Pro Ser Leu Asp Gln Glu Glu Thr Ser Glu Ala Asp Gln  
 340 345 350

Arg Val Val Tyr Lys Phe Gln Ala Asp Gln Asp Lys Phe Asp Leu Asp  
 355 360 365

Arg Asp Pro Glu Gly Val Trp Leu Val Ser Gly Pro Lys Val Glu Arg  
 370 375 380

Leu Tyr Ala Met Thr Asn Phe Asp His Glu Glu Ala Ile Met Arg Phe  
 385 390 395 400

Ser Arg Gln Leu Arg Gly Met Gly Val Asp Gln Ala Leu Arg Asp Lys  
 405 410 415

Gly Ala Gln Ser Gly Asp Leu Val Gln Val Glu Asp Phe Val Phe Glu  
 420 425 430

Phe Met Asp  
 435

<210> 73  
 <211> 1338  
 <212> DNA  
 <213> *Alloiococcus otitidis*

<220>  
 <221> CDS  
 <222> (25)..(1338)  
 <223>

<400> 73  
 aagagaaaga aagaaggtgt actg atg gct aat cct tta gta gcc ata atc 51  
 Met Ala Asn Pro Leu Val Ala Ile Ile  
 1 5  
 ggc cgg cct aat gtc ggc aag tca act att ttc aac cgg att att gga 99  
 Gly Arg Pro Asn Val Gly Lys Ser Thr Ile Phe Asn Arg Ile Ile Gly  
 10 15 20 25  
 gac cgc tta gcc att gtc cag gat gaa ccc ggg gtc acc cgg gac cgt 147  
 Asp Arg Leu Ala Ile Val Gln Asp Glu Pro Gly Val Thr Arg Asp Arg  
 30 35 40  
 att tat gcc gat gct gaa tgg ttg ggc aaa gac ttt tct gtt ata gat 195  
 Ile Tyr Ala Asp Ala Glu Trp Leu Gly Lys Asp Phe Ser Val Ile Asp  
 45 50 55  
 acg gga gga atc act ttt gat gat ttg ccc ttg cat gaa gaa ata aaa 243  
 Thr Gly Gly Ile Thr Phe Asp Asp Leu Pro Leu His Glu Glu Ile Lys  
 60 65 70  
 gtc caa gct gaa att gcc att gat gaa gca gat gtc atc gtc atg gta 291  
 Val Gln Ala Glu Ile Ala Ile Asp Glu Ala Asp Val Ile Val Met Val  
 75 80 85

acc agt gtc aaa gag ggc att aca gac ttg gat gac cag gta gcc tta Thr Ser Val Lys Glu Gly Ile Thr Asp Leu Asp Asp Gln Val Ala Leu 90 95 100 105	339
att ttg cag cag tcc aac aaa ccc gtg gtc ctt gct gtt aat aaa aca Ile Leu Gln Gln Ser Asn Lys Pro Val Val Leu Ala Val Asn Lys Thr 110 115 120	387
gat aat cct gag ctt aga aat gaa ata tat gag ttt tac ggg tta ggc Asp Asn Pro Glu Leu Arg Asn Glu Ile Tyr Glu Phe Tyr Gly Leu Gly 125 130 135	435
ttg ggt gac ccc ctt ccg gta tcc ggg tct cac ggc cta ggc ttt ggg Leu Gly Asp Pro Leu Pro Val Ser Gly Ser His Gly Leu Gly Phe Gly 140 145 150	483
gac ctc tta gac gca gtg gtg gcc aac ttt cct aat gag gcc aat atg Asp Leu Leu Asp Ala Val Val Ala Asn Phe Pro Asn Glu Ala Asn Met 155 160 165	531
gct tat gac caa gat acc att aag ttc tgc ttg att ggt cgt ccc aat Ala Tyr Asp Gln Asp Thr Ile Lys Phe Cys Leu Ile Gly Arg Pro Asn 170 175 180 185	579
gtt ggc aag tct agc cta gtt aat gct att att ggg gaa gac cgg gtt Val Gly Lys Ser Ser Leu Val Asn Ala Ile Ile Gly Glu Asp Arg Val 190 195 200	627
ata gtc tct gaa cta gaa ggg acc acc cgg gat gca att gac act ccc Ile Val Ser Gln Leu Glu Gly Thr Thr Arg Asp Ala Ile Asp Thr Pro 205 210 215	675
ttt atg acc cag gat ggc cag gac tat gtt atg atc gat act gct ggg Phe Met Thr Gln Asp Gly Gln Asp Tyr Val Met Ile Asp Thr Ala Gly 220 225 230	723
atc cgg cgt cgg ggc aag gtc tat gaa aaa act gaa aag tat tct gtt Ile Arg Arg Arg Gly Lys Val Tyr Glu Lys Thr Glu Lys Tyr Ser Val 235 240 245	771
atg cgg gca cag cga gct atc gac cgg tct gat gtg gtc ttg tgt gtc Met Arg Ala Gln Arg Ala Ile Asp Arg Ser Asp Val Val Leu Cys Val 250 255 260 265	819
ctg gat gct gaa aca ggc att aga gac caa gat aag aag gtt ttc ggc Leu Asp Ala Glu Thr Gly Ile Arg Asp Gln Asp Lys Lys Val Phe Gly 270 275 280	867
tat gct cat caa gcc ggc aag gga att att att tta gtc aat aag tgg Tyr Ala His Gln Ala Gly Lys Gly Ile Ile Ile Leu Val Asn Lys Trp 285 290 295	915
gac acg att aaa aaa gag act aac acc atg cga gac ttt gag ttg caa Asp Thr Ile Lys Lys Glu Thr Asn Thr Met Arg Asp Phe Glu Leu Gln 300 305 310	963
att cgc gac caa ttc cgc tac ctc cac tat gcc cca atc ctt ttc gtc	1011

Ile Arg Asp Gln Phe Arg Tyr Leu His Tyr Ala Pro Ile Leu Phe Val  
 315 320 325

tca gcc aag acc aag cag aga ctg gaa gtc atc ccg gaa ttg gtc gac 1059  
 Ser Ala Lys Thr Lys Gln Arg Leu Glu Val Ile Pro Glu Leu Val Asp  
 330 335 340 345

cgg gtc tat tat aac cgc aat caa cgg gtc aag tcc tcc ctc tta aat 1107  
 Arg Val Tyr Tyr Asn Arg Asn Gln Arg Val Lys Ser Ser Leu Leu Asn  
 350 355 360

gat gtg ctg agt gat gca cta gcc agc aat cct gca cct agt aag tca 1155  
 Asp Val Leu Ser Asp Ala Leu Ala Ser Asn Pro Ala Pro Ser Lys Ser  
 365 370 375

ggg aag cga ctc aag gtc ttt tat gcg acc cag gta gcc act aat cca 1203  
 Gly Lys Arg Leu Lys Val Phe Tyr Ala Thr Gln Val Ala Thr Asn Pro  
 380 385 390

cct act ttt gtg gtt ttt gtc aat gat cct gac ctc atg cac ttc tcc 1251  
 Pro Thr Phe Val Val Phe Val Asn Asp Pro Asp Leu Met His Phe Ser  
 395 400 405

tat gag cgc ttt tta gaa aat cga ttc cgc gaa agc ttt gac ttc tat 1299  
 Tyr Glu Arg Phe Leu Glu Asn Arg Phe Arg Glu Ser Phe Asp Phe Tyr  
 410 415 420 425

ggc act ccg att cag ata atc cct aga gca agg aaa taa 1338  
 Gly Thr Pro Ile Gln Ile Ile Pro Arg Ala Arg Lys  
 430 435

&lt;210&gt; 74

&lt;211&gt; 437

&lt;212&gt; PRT

<213> *Alloiococcus otitidis*

&lt;400&gt; 74

Met Ala Asn Pro Leu Val Ala Ile Ile Gly Arg Pro Asn Val Gly Lys  
 1 5 10 15

Ser Thr Ile Phe Asn Arg Ile Ile Gly Asp Arg Leu Ala Ile Val Gln  
 20 25 30

Asp Glu Pro Gly Val Thr Arg Asp Arg Ile Tyr Ala Asp Ala Glu Trp  
 35 40 45

Leu Gly Lys Asp Phe Ser Val Ile Asp Thr Gly Gly Ile Thr Phe Asp  
 50 55 60

Asp Leu Pro Leu His Glu Glu Ile Lys Val Gln Ala Glu Ile Ala Ile  
 65 70 75 80

Asp Glu Ala Asp Val Ile Val Met Val Thr Ser Val Lys Glu Gly Ile  
85 90 95

Thr Asp Leu Asp Asp Gln Val Ala Leu Ile Leu Gln Gln Ser Asn Lys  
100 105 110

Pro Val Val Leu Ala Val Asn Lys Thr Asp Asn Pro Glu Leu Arg Asn  
115 120 125

Glu Ile Tyr Glu Phe Tyr Gly Leu Gly Leu Gly Asp Pro Leu Pro Val  
130 135 140

Ser Gly Ser His Gly Leu Gly Phe Gly Asp Leu Leu Asp Ala Val Val  
145 150 155 160

Ala Asn Phe Pro Asn Glu Ala Asn Met Ala Tyr Asp Gln Asp Thr Ile  
165 170 175

Lys Phe Cys Leu Ile Gly Arg Pro Asn Val Gly Lys Ser Ser Leu Val  
180 185 190

Asn Ala Ile Ile Gly Glu Asp Arg Val Ile Val Ser Glu Leu Glu Gly  
195 200 205

Thr Thr Arg Asp Ala Ile Asp Thr Pro Phe Met Thr Gln Asp Gly Gln  
210 215 220

Asp Tyr Val Met Ile Asp Thr Ala Gly Ile Arg Arg Arg Gly Lys Val  
225 230 235 240

Tyr Glu Lys Thr Glu Lys Tyr Ser Val Met Arg Ala Gln Arg Ala Ile  
245 250 255

Asp Arg Ser Asp Val Val Leu Cys Val Leu Asp Ala Glu Thr Gly Ile  
260 265 270

Arg Asp Gln Asp Lys Lys Val Phe Gly Tyr Ala His Gln Ala Gly Lys  
275 280 285

Gly Ile Ile Ile Leu Val Asn Lys Trp Asp Thr Ile Lys Lys Glu Thr  
290 295 300

Asn Thr Met Arg Asp Phe Glu Leu Gln Ile Arg Asp Gln Phe Arg Tyr  
 305 310 315 320

Leu His Tyr Ala Pro Ile Leu Phe Val Ser Ala Lys Thr Lys Gln Arg  
 325 330 335

Leu Glu Val Ile Pro Glu Leu Val Asp Arg Val Tyr Tyr Asn Arg Asn  
 340 345 350

Gln Arg Val Lys Ser Ser Leu Leu Asn Asp Val Leu Ser Asp Ala Leu  
 355 360 365

Ala Ser Asn Pro Ala Pro Ser Lys Ser Gly Lys Arg Leu Lys Val Phe  
 370 375 380

Tyr Ala Thr Gln Val Ala Thr Asn Pro Pro Thr Phe Val Val Phe Val  
 385 390 395 400

Asn Asp Pro Asp Leu Met His Phe Ser Tyr Glu Arg Phe Leu Glu Asn  
 405 410 415

Arg Phe Arg Glu Ser Phe Asp Phe Tyr Gly Thr Pro Ile Gln Ile Ile  
 420 425 430

Pro Arg Ala Arg Lys  
 435

<210> 75  
 <211> 3324  
 <212> DNA  
 <213> *Alloioococcus otitidis*

<220>  
 <221> CDS  
 <222> (10)..(3324)  
 <223>

<400> 75  
 aataaaaga ttg aaa caa ata tgt ctt aga cga aga ggt gac aag atg act 51  
 Met Lys Gln Ile Cys Leu Arg Arg Arg Gly Asp Lys Met Thr  
 1 5 10

ttt acc cac tta caa gtg acc agt gct tac acc ttg atg gct tcg acc 99  
 Phe Thr His Leu Gln Val Thr Ser Ala Tyr Thr Leu Met Ala Ser Thr  
 15 20 25 30

atc caa ttg ccc ctc ctg atg gac cgc ctg aag gag ctt ggc atg gag 147  
 Ile Gln Leu Pro Leu Leu Met Asp Arg Leu Lys Glu Leu Gly Met Glu



35	40	45	
gct gtt gcc ttg acc gac cac aat gtt atg cat gga gcg gtc gaa ttt Ala Val Ala Leu Thr Asp His Asn Val Met His Gly Ala Val Glu Phe 50 55 60			195
tac caa gaa gcc aaa aag cat ggc att aaa ccc att atg gga cta cgg Tyr Gln Glu Ala Lys Lys His Gly Ile Lys Pro Ile Met Gly Leu Arg 65 70 75			243
gct gac cta gac gaa gga ata acc gtc acc ctc ctg gct aaa aac aag Ala Asp Leu Asp Glu Gly Ile Thr Val Thr Leu Leu Ala Lys Asn Lys 80 85 90			291
gct ggc tac cag gct ctc tta gcc tta tgc act gac ctt caa gtt aac Ala Gly Tyr Gln Ala Leu Leu Ala Leu Ser Thr Asp Leu Gln Val Asn 95 100 105 110			339
aag cag gct att aca ctt gac caa gtc cgt tct gtg gcc cag gac ctc Lys Gln Ala Ile Thr Leu Asp Gln Val Arg Ser Val Ala Gln Asp Leu 115 120 125			387
tat aca ata ttc cca agc tct gac cca aaa gtg aaa gca gac ctc tta Tyr Thr Ile Phe Pro Ser Ser Asp Pro Lys Val Lys Ala Asp Leu Leu 130 135 140			435
gat aag cag gca agc aat ttg acc gcg atg act cag aac ctg ccc cat Asp Lys Gln Ala Ser Asn Leu Thr Ala Met Thr Gln Asn Leu Pro His 145 150 155			483
tca tat ttg ggt ctg gtg cca gac caa gat caa aaa att tac cag tta Ser Tyr Leu Gly Leu Val Pro Asp Gln Asp Gln Lys Ile Tyr Gln Leu 160 165 170			531
gcc cgg acc ttg tca gat tct gga ggt ttg aaa gtc tta gcc tta tct Ala Arg Thr Leu Ser Asp Ser Gly Gly Leu Lys Val Leu Ala Leu Ser 175 180 185 190			579
gac gtc cgt tgc ttg gaa gaa agc caa gtc tcc act ttg gaa atc tta Asp Val Arg Cys Leu Glu Glu Ser Gln Val Ser Thr Leu Glu Ile Leu 195 200 205			627
agc cac atc aaa gcc aac cag aaa att caa ttt gac acc cag gct cgg Ser His Ile Lys Ala Asn Gln Lys Ile Gln Phe Asp Thr Gln Ala Arg 210 215 220			675
gaa aat tat gcc ctg cgc agt ccc caa gaa atg gag tct ttt ttt aac Glu Asn Tyr Ala Leu Arg Ser Pro Gln Glu Met Glu Ser Phe Phe Asn 225 230 235			723
cag gtg ggt tta ggt cag gcc ctt aaa aat act aaa gat gta gcc cag Gln Val Gly Leu Gly Gln Ala Leu Lys Asn Thr Lys Asp Val Ala Gln 240 245 250			771
tca gta gac tgg tcc ctg gac ctc ggt cag gct aaa ttg cct gca ttt Ser Val Asp Trp Ser Leu Asp Leu Gly Gln Ala Lys Leu Pro Ala Phe 255 260 265 270			819

gac ctg ccg gaa ggg gag acc aag gac tcc tac ctt ggc aag ctt gcc Asp Leu Pro Glu Gly Glu Thr Lys Asp Ser Tyr Leu Gly Lys Leu Ala 275 280 285	867
caa aaa gga ctc caa gaa cgg gtt cca ggc tac ggc caa gac tac caa Gln Lys Gly Leu Gln Glu Arg Val Pro Gly Tyr Gly Gln Asp Tyr Gln 290 295 300	915
gac cgt cta gac aag gaa cta gcg gtt att tct tcc atg ggc ttt tcg Asp Arg Leu Asp Lys Glu Leu Ala Val Ile Ser Ser Met Gly Phe Ser 305 310 315	963
gac tac ttc ctg att gtt tgg gac ctg atg caa ttt gcc cgc cag gaa Asp Tyr Phe Leu Ile Val Trp Asp Leu Met Gln Phe Ala Arg Gln Glu 320 325 330	1011
aaa att gag act ggt ttt ggc cgg ggg tca gct gcc gct tct ttg gta Lys Ile Glu Thr Gly Phe Gly Arg Gly Ser Ala Ala Ala Ser Leu Val 335 340 345 350	1059
tct tat gcc ctc tac att acg ggg gta gat ccc atc cat tat gac ctc Ser Tyr Ala Leu Tyr Ile Thr Gly Val Asp Pro Ile His Tyr Asp Leu 355 360 365	1107
ctc ttt gaa cgt ttt ttg aac aag gac cgc ttt acc atg cct gat att Leu Phe Glu Arg Phe Leu Asn Lys Asp Arg Phe Thr Met Pro Asp Ile 370 375 380	1155
gac cta gac ttc cca gac aac aag cgc cag gtc atc ttg gac tat gtc Asp Leu Asp Phe Pro Asp Asn Lys Arg Gln Val Ile Leu Asp Tyr Val 385 390 395	1203
tac cgg aag tat ggt cct gac cat gtg gcc caa att ttg acc ttt ggg Tyr Arg Lys Tyr Gly Pro Asp His Val Ala Gln Ile Leu Thr Phe Gly 400 405 410	1251
acc ttt gcg gct aag tcc tcc atc agg gaa att atg cgg acc ttg ggt Thr Phe Ala Ala Lys Ser Ser Ile Arg Glu Ile Met Arg Thr Leu Gly 415 420 425 430	1299
tac aag aat gaa gac atg aag acc tgg tcc cag gcc ata cca gat acc Tyr Lys Asn Glu Asp Met Lys Thr Trp Ser Gln Ala Ile Pro Asp Thr 435 440 445	1347
gtc aac atc agc ttg tca aag gcc tat gac gag tcg aaa gac ctt caa Val Asn Ile Ser Leu Ser Lys Ala Tyr Asp Glu Ser Lys Asp Leu Gln 450 455 460	1395
aaa ctg gtc cag caa agc cat gaa aat gag cgg atc ttt gcc atg gcc Lys Leu Val Gln Gln Ser His Glu Asn Glu Arg Ile Phe Ala Met Ala 465 470 475	1443
cag gat atc gaa ggc ctg cca agg aac tat tca acc cat gcg gcc ggt Gln Asp Ile Glu Gly Leu Pro Arg Asn Tyr Ser Thr His Ala Ala Gly 480 485 490	1491

gtc gtc atg tca gac cag ccc cta atc cat tcc ctt ccc cta caa gat Val Val Met Ser Asp Gln Pro Leu Ile His Ser Leu Pro Leu Gln Asp 495 500 505 510	1539
ggc aac gga aag gtc ccc aac acc caa ttt acc atg gag gat gtt gaa Gly Asn Gly Lys Val Pro Asn Thr Gln Phe Thr Met Glu Asp Val Glu 515 520 525	1587
gcg gtc ggc tta ctc aag atg gac ttt ttg agt tta aaa aat tta acc Ala Val Gly Leu Leu Lys Met Asp Phe Leu Ser Leu Lys Asn Leu Thr 530 535 540	1635
atc cta gca gac tgc ttg aac ttt agc cag tat gaa ggg cag gga ggg Ile Leu Ala Asp Cys Leu Asn Phe Ser Gln Tyr Glu Gly Gln Gly Gly 545 550 555	1683
ggc ata agt aaa caa gat ata cca atc gac gac cct aag acc ctg gat Gly Ile Ser Lys Gln Asp Ile Pro Ile Asp Asp Pro Lys Thr Leu Asp 560 565 570	1731
ctt ttt gcc cgg gga gac aca aat ggg gtc ttc caa ttt gaa aaa gag Leu Phe Ala Arg Gly Asp Thr Asn Gly Val Phe Gln Phe Glu Lys Glu 575 580 585 590	1779
gga atc aaa aaa gtc ctc cgc cag ctt caa ccc act tct ttt gaa gat Gly Ile Lys Lys Val Leu Arg Gln Leu Gln Pro Thr Ser Phe Glu Asp 595 600 605	1827
atc gtc gcc acc aac gcc ctc tac cgc ccc ggt ccc atg ggg caa att Ile Val Ala Thr Asn Ala Leu Tyr Arg Pro Gly Pro Met Gly Gln Ile 610 615 620	1875
gag aat tat att aac cgt aaa cat ggt caa gaa aaa att atc tac ccc Glu Asn Tyr Ile Asn Arg Lys His Gly Gln Glu Lys Ile Ile Tyr Pro 625 630 635	1923
cat gaa gac tta aag gac atc ctt gaa gtc act tat ggc att att gtc His Glu Asp Leu Lys Asp Ile Leu Glu Val Thr Tyr Gly Ile Ile Val 640 645 650	1971
tac cag gaa caa gtc atg cag gta gct acc caa cta gct ggc tat agt Tyr Gln Glu Gln Val Met Gln Val Ala Thr Gln Leu Ala Gly Tyr Ser 655 660 665 670	2019
ttg tcg gaa gct gac caa ttg cgg cgg act atg tcc aaa aaa atc cag Leu Ser Glu Ala Asp Gln Leu Arg Arg Thr Met Ser Lys Lys Ile Gln 675 680 685	2067
tca gaa atg gac cag gga cgg gaa aaa ttt ata aga gga gcc ttg gac Ser Glu Met Asp Gln Gly Arg Glu Lys Phe Ile Arg Gly Ala Leu Asp 690 695 700	2115
aag ggc tac agt gag tca gta gcc cga gag gtt tat aac tat att gca Lys Gly Tyr Ser Glu Ser Val Ala Arg Glu Val Tyr Asn Tyr Ile Ala 705 710 715	2163
aag ttt gct aac tac ggc ttt aac cgt gcc cat gct gtt gcc tac tcc	2211

Lys Phe Ala Asn Tyr Gly Phe Asn Arg Ala His Ala Val Ala Tyr Ser	
720 725 730	
atg ctt gcc tac cat atg gcc tac ttt aag gtc cac cag cct aaa tct	2259
Met Leu Ala Tyr His Met Ala Tyr Phe Lys Val His Gln Pro Lys Ser	
735 740 745 750	
ttt ttt gcg gct gtg atg aag gca gac tgg ggt aac aag gct aaa att	2307
Phe Phe Ala Ala Val Met Lys Ala Asp Trp Gly Asn Lys Ala Lys Ile	
755 760 765	
tac aag tat gcc cat gaa gtc cgg gct aga aaa att aaa cta cta aaa	2355
Tyr Lys Tyr Ala His Glu Val Arg Ala Arg Lys Ile Lys Leu Leu Lys	
770 775 780	
cca gac atc aac caa agc ctt gga tct ttt acg gtt cgg cag aat ggc	2403
Pro Asp Ile Asn Gln Ser Leu Gly Ser Phe Thr Val Arg Gln Asn Gly	
785 790 795	
att caa gtg ggg ctt aag atg gtc aag ggg gtg gct agc ccc ttt gtc	2451
Ile Gln Val Gly Leu Lys Met Val Lys Gly Val Ala Ser Pro Phe Val	
800 805 810	
aac cac atc ctt gaa att cgg aaa gaa aag gga gct ttt acc agc ctg	2499
Asn His Ile Leu Glu Ile Arg Lys Glu Lys Gly Ala Phe Thr Ser Leu	
815 820 825 830	
cgt gac ttt tgt gaa aaa att gac agc caa ttc tta agt caa gac ccc	2547
Arg Asp Phe Cys Glu Lys Ile Asp Ser Gln Phe Leu Ser Gln Asp Pro	
835 840 845	
att gaa gca ttg att ttg gtg ggg gcc ttt gac caa atg ggc cct aat	2595
Ile Glu Ala Leu Ile Leu Val Gly Ala Phe Asp Gln Met Gly Pro Asn	
850 855 860	
cgg cgg acc atg tta gcg ggc ttg gaa gca acg att gaa ttc gtg gcc	2643
Arg Arg Thr Met Leu Ala Gly Leu Glu Ala Thr Ile Glu Phe Val Ala	
865 870 875	
aaa agt tgc ggc aat atc acc ctt ttt gac act ctc aag ccc cgc caa	2691
Lys Ser Ser Gly Asn Ile Thr Leu Phe Asp Thr Leu Lys Pro Arg Gln	
880 885 890	
gaa gac ctg gaa gag ttt agc cca aag gac ctc att caa tat gaa gaa	2739
Glu Asp Leu Glu Glu Phe Ser Pro Lys Asp Leu Ile Gln Tyr Glu Glu	
895 900 905 910	
gaa tta acc ggt ttt tac ttc tcc agc cac ccc ttg agc cgg tat gac	2787
Glu Leu Thr Gly Phe Tyr Phe Ser Ser His Pro Leu Ser Arg Tyr Asp	
915 920 925	
tcc ctg cga cag gac tta aaa acg tcc ttt ata gct gat tta gaa gag	2835
Ser Leu Arg Gln Asp Leu Lys Thr Ser Phe Ile Ala Asp Leu Glu Glu	
930 935 940	
ggc caa tct tgc caa gtt tta ggt cag ctg gtt caa gtc cgg aaa act	2883
Gly Gln Ser Cys Gln Val Leu Gly Gln Leu Val Gln Val Arg Lys Thr	

945	950	955	
cag act aga aac caa caa ccc atg gcc ttt gtt agc ctg gct gac caa			2931
Gln Thr Arg Asn Gln Gln Pro Met Ala Phe Val Ser Leu Ala Asp Gln			
960	965	970	
aca gga caa att agc ctg gtg gtc ttt ccg aat gta tac cgc gaa tgc			2979
Thr Gly Gln Ile Ser Leu Val Val Phe Pro Asn Val Tyr Arg Glu Cys			
975	980	985	990
cta cct tac ctc aaa gaa gga gtg gtc ctg gtc gtc tca ggc aag gta			3027
Leu Pro Tyr Leu Lys Glu Gly Val Val Leu Val Val Ser Gly Lys Val			
995	1000		1005
gaa gtt agg aag gga gaa atc cag cta aaa gtc cag acc atg aaa gag			3075
Glu Val Arg Lys Gly Glu Ile Gln Leu Lys Val Gln Thr Met Lys Glu			
1010	1015		1020
gcc agc cag gtc caa aaa gag act aag cag ctt tac ctg aaa ttt gct			3123
Ala Ser Gln Val Gln Lys Glu Thr Lys Gln Leu Tyr Leu Lys Phe Ala			
1025	1030		1035
gac ttg aac caa gat aaa gaa agt ttt cgt caa gtg caa aag atc ttg			3171
Asp Leu Asn Gln Asp Lys Glu Ser Phe Arg Gln Val Gln Lys Ile Leu			
1040	1045		1050
gcc cga cat ccc ggc cag aag cga gtg att gtt tac gac cag gcc agc			3219
Ala Arg His Pro Gly Gln Lys Arg Val Ile Val Tyr Asp Gln Ala Ser			
1055	1060		1065
cag caa gca ctc cag ctc aaa gca aaa ttt aat ttc gac gga cgg acg			3267
Gln Gln Ala Leu Gln Leu Lys Ala Lys Phe Asn Phe Asp Gly Arg Thr			
1075	1080		1085
gat acc cta aac cag ctc cag gac ctc cta ggc cag gat tct tgt atc			3315
Asp Thr Leu Asn Gln Leu Gln Asp Leu Leu Gly Gln Asp Ser Cys Ile			
1090	1095		1100
tta aaa taa			3324
Leu Lys			

&lt;210&gt; 76

&lt;211&gt; 1104

&lt;212&gt; PRT

<213> *Alloiococcus otitidis*

&lt;400&gt; 76

Met	Lys	Gln	Ile	Cys	Leu	Arg	Arg	Arg	Gly	Asp	Lys	Met	Thr	Phe	Thr
1				5					10					15	

His	Leu	Gln	Val	Thr	Ser	Ala	Tyr	Thr	Leu	Met	Ala	Ser	Thr	Ile	Gln
		20						25					30		

Leu Pro Leu Leu Met Asp Arg Leu Lys Glu Leu Gly Met Glu Ala Val

35

40

45

Ala Leu Thr Asp His Asn Val Met His Gly Ala Val Glu Phe Tyr Gln  
 50 55 60

Glu Ala Lys Lys His Gly Ile Lys Pro Ile Met Gly Leu Arg Ala Asp  
 65 70 75 80

Leu Asp Glu Gly Ile Thr Val Thr Leu Leu Ala Lys Asn Lys Ala Gly  
 85 90 95

Tyr Gln Ala Leu Leu Ala Leu Ser Thr Asp Leu Gln Val Asn Lys Gln  
 100 105 110

Ala Ile Thr Leu Asp Gln Val Arg Ser Val Ala Gln Asp Leu Tyr Thr  
 115 120 125

Ile Phe Pro Ser Ser Asp Pro Lys Val Lys Ala Asp Leu Leu Asp Lys  
 130 135 140

Gln Ala Ser Asn Leu Thr Ala Met Thr Gln Asn Leu Pro His Ser Tyr  
 145 150 155 160

Leu Gly Leu Val Pro Asp Gln Asp Gln Lys Ile Tyr Gln Leu Ala Arg  
 165 170 175

Thr Leu Ser Asp Ser Gly Gly Leu Lys Val Leu Ala Leu Ser Asp Val  
 180 185 190

Arg Cys Leu Glu Glu Ser Gln Val Ser Thr Leu Glu Ile Leu Ser His  
 195 200 205

Ile Lys Ala Asn Gln Lys Ile Gln Phe Asp Thr Gln Ala Arg Glu Asn  
 210 215 220

Tyr Ala Leu Arg Ser Pro Gln Glu Met Glu Ser Phe Phe Asn Gln Val  
 225 230 235 240

Gly Leu Gly Gln Ala Leu Lys Asn Thr Lys Asp Val Ala Gln Ser Val  
 245 250 255

Asp Trp Ser Leu Asp Leu Gly Gln Ala Lys Leu Pro Ala Phe Asp Leu  
 260 265 270

Pro Glu Gly Glu Thr Lys Asp Ser Tyr Leu Gly Lys Leu Ala Gln Lys  
 275 280 285

Gly Leu Gln Glu Arg Val Pro Gly Tyr Gly Gln Asp Tyr Gln Asp Arg  
 290 295 300

Leu Asp Lys Glu Leu Ala Val Ile Ser Ser Met Gly Phe Ser Asp Tyr  
 305 310 315 320

Phe Leu Ile Val Trp Asp Leu Met Gln Phe Ala Arg Gln Glu Lys Ile  
 325 330 335

Glu Thr Gly Phe Gly Arg Gly Ser Ala Ala Ala Ser Leu Val Ser Tyr  
 340 345 350

Ala Leu Tyr Ile Thr Gly Val Asp Pro Ile His Tyr Asp Leu Leu Phe  
 355 360 365

Glu Arg Phe Leu Asn Lys Asp Arg Phe Thr Met Pro Asp Ile Asp Leu  
 370 375 380

Asp Phe Pro Asp Asn Lys Arg Gln Val Ile Leu Asp Tyr Val Tyr Arg  
 385 390 395 400

Lys Tyr Gly Pro Asp His Val Ala Gln Ile Leu Thr Phe Gly Thr Phe  
 405 410 415

Ala Ala Lys Ser Ser Ile Arg Glu Ile Met Arg Thr Leu Gly Tyr Lys  
 420 425 430

Asn Glu Asp Met Lys Thr Trp Ser Gln Ala Ile Pro Asp Thr Val Asn  
 435 440 445

Ile Ser Leu Ser Lys Ala Tyr Asp Glu Ser Lys Asp Leu Gln Lys Leu  
 450 455 460

Val Gln Gln Ser His Glu Asn Glu Arg Ile Phe Ala Met Ala Gln Asp  
 465 470 475 480

Ile Glu Gly Leu Pro Arg Asn Tyr Ser Thr His Ala Ala Gly Val Val  
 485 490 495

Met Ser Asp Gln Pro Leu Ile His Ser Leu Pro Leu Gln Asp Gly Asn  
500 505 510

Gly Lys Val Pro Asn Thr Gln Phe Thr Met Glu Asp Val Glu Ala Val  
515 520 525

Gly Leu Leu Lys Met Asp Phe Leu Ser Leu Lys Asn Leu Thr Ile Leu  
530 535 540

Ala Asp Cys Leu Asn Phe Ser Gln Tyr Glu Gly Gln Gly Gly Gly Ile  
545 550 555 560

Ser Lys Gln Asp Ile Pro Ile Asp Asp Pro Lys Thr Leu Asp Leu Phe  
565 570 575

Ala Arg Gly Asp Thr Asn Gly Val Phe Gln Phe Glu Lys Glu Gly Ile  
580 585 590

Lys Lys Val Leu Arg Gln Leu Gln Pro Thr Ser Phe Glu Asp Ile Val  
595 600 605

Ala Thr Asn Ala Leu Tyr Arg Pro Gly Pro Met Gly Gln Ile Glu Asn  
610 615 620

Tyr Ile Asn Arg Lys His Gly Gln Glu Lys Ile Ile Tyr Pro His Glu  
625 630 635 640

Asp Leu Lys Asp Ile Leu Glu Val Thr Tyr Gly Ile Ile Val Tyr Gln  
645 650 655

Glu Gln Val Met Gln Val Ala Thr Gln Leu Ala Gly Tyr Ser Leu Ser  
660 665 670

Glu Ala Asp Gln Leu Arg Arg Thr Met Ser Lys Lys Ile Gln Ser Glu  
675 680 685

Met Asp Gln Gly Arg Glu Lys Phe Ile Arg Gly Ala Leu Asp Lys Gly  
690 695 700

Tyr Ser Glu Ser Val Ala Arg Glu Val Tyr Asn Tyr Ile Ala Lys Phe  
705 710 715 720



Ala Asn Tyr Gly Phe Asn Arg Ala His Ala Val Ala Tyr Ser Met Leu  
725 730 735

Ala Tyr His Met Ala Tyr Phe Lys Val His Gln Pro Lys Ser Phe Phe  
740 745 750

Ala Ala Val Met Lys Ala Asp Trp Gly Asn Lys Ala Lys Ile Tyr Lys  
755 760 765

Tyr Ala His Glu Val Arg Ala Arg Lys Ile Lys Leu Leu Lys Pro Asp  
770 775 780

Ile Asn Gln Ser Leu Gly Ser Phe Thr Val Arg Gln Asn Gly Ile Gln  
785 790 795 800

Val Gly Leu Lys Met Val Lys Gly Val Ala Ser Pro Phe Val Asn His  
805 810 815

Ile Leu Glu Ile Arg Lys Glu Lys Gly Ala Phe Thr Ser Leu Arg Asp  
820 825 830

Phe Cys Glu Lys Ile Asp Ser Gln Phe Leu Ser Gln Asp Pro Ile Glu  
835 840 845

Ala Leu Ile Leu Val Gly Ala Phe Asp Gln Met Gly Pro Asn Arg Arg  
850 855 860

Thr Met Leu Ala Gly Leu Glu Ala Thr Ile Glu Phe Val Ala Lys Ser  
865 870 875 880

Ser Gly Asn Ile Thr Leu Phe Asp Thr Leu Lys Pro Arg Gln Glu Asp  
885 890 895

Leu Glu Glu Phe Ser Pro Lys Asp Leu Ile Gln Tyr Glu Glu Glu Leu  
900 905 910

Thr Gly Phe Tyr Phe Ser Ser His Pro Leu Ser Arg Tyr Asp Ser Leu  
915 920 925

Arg Gln Asp Leu Lys Thr Ser Phe Ile Ala Asp Leu Glu Glu Gly Gln  
930 935 940

Ser Cys Gln Val Leu Gly Gln Leu Val Gln Val Arg Lys Thr Gln Thr

945                      950                      955                      960  
 Arg Asn Gln Gln Pro Met Ala Phe Val Ser Leu Ala Asp Gln Thr Gly  
                                  965                      970                      975  
 Gln Ile Ser Leu Val Val Phe Pro Asn Val Tyr Arg Glu Cys Leu Pro  
                                  980                      985                      990  
 Tyr Leu Lys Glu Gly Val Val Leu Val Val Ser Gly Lys Val Glu Val  
                                  995                      1000                      1005  
 Arg Lys Gly Glu Ile Gln Leu Lys Val Gln Thr Met Lys Glu Ala Ser  
                                  1010                      1015                      1020  
 Gln Val Gln Lys Glu Thr Lys Gln Leu Tyr Leu Lys Phe Ala Asp Leu  
                                  1025                      1030                      1035                      1040  
 Asn Gln Asp Lys Glu Ser Phe Arg Gln Val Gln Lys Ile Leu Ala Arg  
                                  1045                      1050                      1055  
 His Pro Gly Gln Lys Arg Val Ile Val Tyr Asp Gln Ala Ser Gln Gln  
                                  1060                      1065                      1070  
 Ala Leu Gln Leu Lys Ala Lys Phe Asn Phe Asp Gly Arg Thr Asp Thr  
                                  1075                      1080                      1085  
 Leu Asn Gln Leu Gln Asp Leu Leu Gly Gln Asp Ser Cys Ile Leu Lys  
                                  1090                      1095                      1100

<210> 77  
 <211> 1212  
 <212> DNA  
 <213> *Alloioococcus otitidis*

<220>  
 <221> CDS  
 <222> (7) .. (1212)  
 <223>

<400> 77  
 acaaag atg ctg aaa aat aaa aag ata gcc tta tat gtt act ggt ggt  
 Met Leu Lys Asn Lys Lys Ile Ala Leu Tyr Val Thr Gly Gly  
                                  1                      5                      10

ata gca gta tac aaa tca ctt tac tta ctt agg gaa atc atc aaa caa Ile Ala Val Tyr Lys Ser Leu Tyr Leu Leu Arg Glu Ile Ile Lys Gln 15 20 25 30	96
ggc ggg gag gtc cgg gtt gcc atg act caa gca gct tgt caa ttt gtt Gly Gly Glu Val Arg Val Ala Met Thr Gln Ala Ala Cys Gln Phe Val 35 40 45	144
aac ccc tta tct ttt cag gtt tta agc caa aaa aag gtt cag att gac Asn Pro Leu Ser Phe Gln Val Leu Ser Gln Lys Lys Val Gln Ile Asp 50 55 60	192
act ttt gaa gaa ggt cag ccc gaa tcg gtc agt cac att gat ttg acg Thr Phe Glu Glu Gly Gln Pro Glu Ser Val Ser His Ile Asp Leu Thr 65 70 75	240
gat tgg gcc gac tac tcc atc gtg gct ccg gca act gcc aat atc atc Asp Trp Ala Asp Tyr Ser Ile Val Ala Pro Ala Thr Ala Asn Ile Ile 80 85 90	288
ggc aag ctg gcc aat ggg att ggg gac gat ttt gtt tca aca gcc ttg Gly Lys Leu Ala Asn Gly Ile Gly Asp Asp Phe Val Ser Thr Ala Leu 95 100 105 110	336
ttg gca acg gac cac ccc att ttt tta gtc cca gcc atg aac acc aag Leu Ala Thr Asp His Pro Ile Phe Leu Val Pro Ala Met Asn Thr Lys 115 120 125	384
atg tat gaa aat ccc gct ctt aag aaa aac aag gcc ttc ctt att gaa Met Tyr Glu Asn Pro Ala Leu Lys Lys Asn Lys Ala Phe Leu Ile Glu 130 135 140	432
cag ggc cat tac tgg atg gag ccg gat att gga ttt tta gca gag ggc Gln Gly His Tyr Trp Met Glu Pro Asp Ile Gly Phe Leu Ala Glu Gly 145 150 155	480
tac gaa ggc ttg ggt cgt ttt cca gac cta gac cgg att atg gcg gaa Tyr Glu Gly Leu Gly Arg Phe Pro Asp Leu Asp Arg Ile Met Ala Glu 160 165 170	528
ttt aac cat ttt att att gct agg aat cca ggt atc cta tca gga aaa Phe Asn His Phe Ile Ile Ala Arg Asn Pro Gly Ile Leu Ser Gly Lys 175 180 185 190	576
aaa gtc ctc gtc aca gca ggt ggg acg gtg gag cgg att gat ccc gtc Lys Val Leu Val Thr Ala Gly Gly Thr Val Glu Arg Ile Asp Pro Val 195 200 205	624
cgg tat att tcc aat gat tct tct ggt aag atg ggc cac caa ctt gct Arg Tyr Ile Ser Asn Asp Ser Ser Gly Lys Met Gly His Gln Leu Ala 210 215 220	672
caa gcg gcc tat gaa gct ggg gcc cag gtt agc ttg gta aca gcc agt Gln Ala Ala Tyr Glu Ala Gly Ala Gln Val Ser Leu Val Thr Ala Ser 225 230 235	720
gac ttg ccg acc agt ccc ttt att gac cgc ttt cag gtg gag tcc acc	768

Asp	Leu	Pro	Thr	Ser	Pro	Phe	Ile	Asp	Arg	Phe	Gln	Val	Glu	Ser	Thr		
240						245					250						
tta	gac	ttg	tac	caa	aca	gtt	agt	gac	ctc	tat	gac	cac	cat	gac	att		816
Leu	Asp	Leu	Tyr	Gln	Thr	Val	Ser	Asp	Leu	Tyr	Asp	His	His	Asp	Ile		
255				260					265						270		
ctc	atg	atg	gcc	gca	gcg	gtg	tct	gac	tac	cgg	cca	gtc	aac	cgg	tca		864
Leu	Met	Met	Ala	Ala	Ala	Val	Ser	Asp	Tyr	Arg	Pro	Val	Asn	Arg	Ser		
			275					280					285				
gac	aaa	aag	atg	aaa	aag	caa	gat	aat	tta	acc	att	gaa	ctg	gaa	aaa		912
Asp	Lys	Lys	Met	Lys	Lys	Gln	Asp	Asn	Leu	Thr	Ile	Glu	Leu	Glu	Lys		
			290					295				300					
aat	cct	gat	att	ttg	gcc	gaa	atg	ggc	cgg	cgg	aaa	gac	caa	caa	atc		960
Asn	Pro	Asp	Ile	Leu	Ala	Glu	Met	Gly	Arg	Arg	Lys	Asp	Gln	Gln	Ile		
	305						310					315					
aat	gtc	ggc	ttt	gca	gca	gaa	acc	cat	aac	ctt	gaa	gaa	tat	gcc	caa		1008
Asn	Val	Gly	Phe	Ala	Ala	Glu	Thr	His	Asn	Leu	Glu	Glu	Tyr	Ala	Gln		
	320					325					330						
aaa	aaa	tta	gcc	tcc	aaa	caa	gct	gac	ttg	atc	gta	gcc	aat	gaa	gtg		1056
Lys	Lys	Leu	Ala	Ser	Lys	Gln	Ala	Asp	Leu	Ile	Val	Ala	Asn	Glu	Val		
335				340					345					350			
ggc	cgg	gga	gac	cgg	ggc	ttt	aat	gcg	gat	gaa	aat	gcg	gcc	ctt	gtt		1104
Gly	Arg	Gly	Asp	Arg	Gly	Phe	Asn	Ala	Asp	Glu	Asn	Ala	Ala	Leu	Val		
			355					360					365				
ttt	tcc	agt	gac	caa	gat	ccg	ctt	gag	ctt	ccc	ctt	cag	tct	aaa	aaa		1152
Phe	Ser	Ser	Asp	Gln	Asp	Pro	Leu	Glu	Leu	Pro	Leu	Gln	Ser	Lys	Lys		
			370					375					380				
gat	atg	gca	aaa	aag	att	att	gaa	gtg	gtg	gcc	agt	aaa	ttg	cct	gct		1200
Asp	Met	Ala	Lys	Lys	Ile	Ile	Glu	Val	Val	Ala	Ser	Lys	Leu	Pro	Ala		
		385					390					395					
tct	ccc	aaa	taa														1212
Ser	Pro	Lys															
		400															
<210>	78																
<211>	401																
<212>	PRT																
<213>	Alloiococcus otitidis																
<400>	78																
Met	Leu	Lys	Asn	Lys	Lys	Ile	Ala	Leu	Tyr	Val	Thr	Gly	Gly	Ile	Ala		
1				5					10					15			
Val	Tyr	Lys	Ser	Leu	Tyr	Leu	Leu	Arg	Glu	Ile	Ile	Lys	Gln	Gly	Gly		
		20						25					30				

Glu Val Arg Val Ala Met Thr Gln Ala Ala Cys Gln Phe Val Asn Pro  
 35 40 45

Leu Ser Phe Gln Val Leu Ser Gln Lys Lys Val Gln Ile Asp Thr Phe  
 50 55 60

Glu Glu Gly Gln Pro Glu Ser Val Ser His Ile Asp Leu Thr Asp Trp  
 65 70 75 80

Ala Asp Tyr Ser Ile Val Ala Pro Ala Thr Ala Asn Ile Ile Gly Lys  
 85 90 95

Leu Ala Asn Gly Ile Gly Asp Asp Phe Val Ser Thr Ala Leu Leu Ala  
 100 105 110

Thr Asp His Pro Ile Phe Leu Val Pro Ala Met Asn Thr Lys Met Tyr  
 115 120 125

Glu Asn Pro Ala Leu Lys Lys Asn Lys Ala Phe Leu Ile Glu Gln Gly  
 130 135 140

His Tyr Trp Met Glu Pro Asp Ile Gly Phe Leu Ala Glu Gly Tyr Glu  
 145 150 155 160

Gly Leu Gly Arg Phe Pro Asp Leu Asp Arg Ile Met Ala Glu Phe Asn  
 165 170 175

His Phe Ile Ile Ala Arg Asn Pro Gly Ile Leu Ser Gly Lys Lys Val  
 180 185 190

Leu Val Thr Ala Gly Gly Thr Val Glu Arg Ile Asp Pro Val Arg Tyr  
 195 200 205

Ile Ser Asn Asp Ser Ser Gly Lys Met Gly His Gln Leu Ala Gln Ala  
 210 215 220

Ala Tyr Glu Ala Gly Ala Gln Val Ser Leu Val Thr Ala Ser Asp Leu  
 225 230 235 240

Pro Thr Ser Pro Phe Ile Asp Arg Phe Gln Val Glu Ser Thr Leu Asp  
 245 250 255

Leu Tyr Gln Thr Val Ser Asp Leu Tyr Asp His His Asp Ile Leu Met  
                   260                  265                  270

Met Ala Ala Ala Val Ser Asp Tyr Arg Pro Val Asn Arg Ser Asp Lys  
                   275                  280                  285

Lys Met Lys Lys Gln Asp Asn Leu Thr Ile Glu Leu Glu Lys Asn Pro  
                   290                  295                  300

Asp Ile Leu Ala Glu Met Gly Arg Arg Lys Asp Gln Gln Ile Asn Val  
                   305                  310                  315                  320

Gly Phe Ala Ala Glu Thr His Asn Leu Glu Glu Tyr Ala Gln Lys Lys  
                   325                  330                  335

Leu Ala Ser Lys Gln Ala Asp Leu Ile Val Ala Asn Glu Val Gly Arg  
                   340                  345                  350

Gly Asp Arg Gly Phe Asn Ala Asp Glu Asn Ala Ala Leu Val Phe Ser  
                   355                  360                  365

Ser Asp Gln Asp Pro Leu Glu Leu Pro Leu Gln Ser Lys Lys Asp Met  
                   370                  375                  380

Ala Lys Lys Ile Ile Glu Val Val Ala Ser Lys Leu Pro Ala Ser Pro  
                   385                  390                  395                  400

Lys

<210> 79  
 <211> 1053  
 <212> DNA  
 <213> Alloiococcus otitidis

<220>  
 <221> CDS  
 <222> (22)..(1053)  
 <223>

<400> 79  
 aagaagaagg gaggaagact g atg aaa att gaa gac caa ctc aaa aaa att  
                                   Met Lys Ile Glu Asp Gln Leu Lys Lys Ile  
                                   1                  5                  10

51

aaa gac caa gac ttg tct ccc ctc tac ctg gtc cag gga gat gac cag  
 Lys Asp Gln Asp Leu Ser Pro Leu Tyr Leu Val Gln Gly Asp Asp Gln  
                   15                  20                  25

99

tac ttg tta gac cag gtt aaa aaa agt ttg agc cag gcc ctt ttg gac Tyr Leu Leu Asp Gln Val Lys Lys Ser Leu Ser Gln Ala Leu Leu Asp 30 35 40	147
cag gat gaa gct tct atg aat ttt ggt caa ttt aat atg atg gct gat Gln Asp Glu Ala Ser Met Asn Phe Gly Gln Phe Asn Met Met Ala Asp 45 50 55	195
agc cta gac atg gcc ttg tct gat gcg gaa tcc tat ccc ttt ttt ggg Ser Leu Asp Met Ala Leu Ser Asp Ala Glu Ser Tyr Pro Phe Phe Gly 60 65 70	243
gac aag cgc ctg gtt tac atc caa gac ccc ttt ttc cta aca ggg gag Asp Lys Arg Leu Val Tyr Ile Gln Asp Pro Phe Phe Leu Thr Gly Glu 75 80 85 90	291
aag cgg aaa aca gat ctg gac cat gac ttg gat cgc ttg ctg gct tac Lys Arg Lys Thr Asp Leu Asp His Asp Leu Asp Arg Leu Leu Ala Tyr 95 100 105	339
ctc caa aac cca gcc gac ttt act gtt ctc gtc ttc ttt gcc ccc tat Leu Gln Asn Pro Ala Asp Phe Thr Val Leu Val Phe Phe Ala Pro Tyr 110 115 120	387
gag aaa ctg gac aag cgg aag aag gtc acc aaa gcc cta ttg cag gaa Glu Lys Leu Asp Lys Arg Lys Lys Val Thr Lys Ala Leu Leu Gln Glu 125 130 135	435
gct gag att ata gat gcc agt tcc cca gac caa aga gat cta aaa gat Ala Glu Ile Ile Asp Ala Ser Ser Pro Asp Gln Arg Asp Leu Lys Asp 140 145 150	483
atg gtc cag aaa aaa gta aag gct cga ggc tac cag ttt gac aaa gga Met Val Gln Lys Lys Val Lys Ala Arg Gly Tyr Gln Phe Asp Lys Gly 155 160 165 170	531
gct tta aag gcc ctg gtt gaa aaa acc aat gcc aac tta agt cgg gtc Ala Leu Lys Ala Leu Val Glu Lys Thr Asn Ala Asn Leu Ser Arg Val 175 180 185	579
atg caa gag ttg gac aag tta ttc ttg tac cat tta gat gac aaa atc Met Gln Glu Leu Asp Lys Leu Phe Leu Tyr His Leu Asp Asp Lys Ile 190 195 200	627
atc acc gtc cag tca gtt gac cag gtc gta tca cca agc ctg gaa agt Ile Thr Val Gln Ser Val Asp Gln Val Val Ser Pro Ser Leu Glu Ser 205 210 215	675
aat gtc ttt agt att aac gac tat att tta agc ggg caa agc cag gct Asn Val Phe Ser Ile Asn Asp Tyr Ile Leu Ser Gly Gln Ser Gln Ala 220 225 230	723
gct ata cgg gcc ttt aat gac tta att caa caa aag gaa gag cca att Ala Ile Arg Ala Phe Asn Asp Leu Ile Gln Gln Lys Glu Glu Pro Ile 235 240 245 250	771

aaa atc atc gcc att atg atg aac caa ttc cgt tta tta ttg cag gtt 819  
 Lys Ile Ile Ala Ile Met Met Asn Gln Phe Arg Leu Leu Leu Gln Val  
 255 260 265  
 aaa ata ttg cgg act aag ggc tac caa caa gga gag atc gct aaa atc 867  
 Lys Ile Leu Arg Thr Lys Gly Tyr Gln Gln Gly Glu Ile Ala Lys Ile  
 270 275 280  
 tta aaa gtt cac ccc tac cgg gtt aag cta gcc ata gag aaa cag gag 915  
 Leu Lys Val His Pro Tyr Arg Val Lys Leu Ala Ile Glu Lys Gln Glu  
 285 290 295  
 att ttt tcc aag caa agt cta tcg acc gcc tac cgc tac tta att gag 963  
 Ile Phe Ser Lys Gln Ser Leu Ser Thr Ala Tyr Arg Tyr Leu Ile Glu  
 300 305 310  
 tca gat cat ttg att aaa acg ggc aag gtg acc tcg caa ttg caa ttt 1011  
 Ser Asp His Leu Ile Lys Thr Gly Lys Val Thr Ser Gln Leu Gln Phe  
 315 320 325 330  
 gaa ctt ttt gcc cta caa ttt aaa gat tct gtc atg aat taa 1053  
 Glu Leu Phe Ala Leu Gln Phe Lys Asp Ser Val Met Asn  
 335 340

&lt;210&gt; 80

&lt;211&gt; 343

&lt;212&gt; PRT

<213> *Alloiococcus otitidis*

&lt;400&gt; 80

Met Lys Ile Glu Asp Gln Leu Lys Lys Ile Lys Asp Gln Asp Leu Ser  
 1 5 10 15

Pro Leu Tyr Leu Val Gln Gly Asp Asp Gln Tyr Leu Leu Asp Gln Val  
 20 25 30

Lys Lys Ser Leu Ser Gln Ala Leu Leu Asp Gln Asp Glu Ala Ser Met  
 35 40 45

Asn Phe Gly Gln Phe Asn Met Met Ala Asp Ser Leu Asp Met Ala Leu  
 50 55 60

Ser Asp Ala Glu Ser Tyr Pro Phe Phe Gly Asp Lys Arg Leu Val Tyr  
 65 70 75 80

Ile Gln Asp Pro Phe Phe Leu Thr Gly Glu Lys Arg Lys Thr Asp Leu  
 85 90 95

Asp His Asp Leu Asp Arg Leu Leu Ala Tyr Leu Gln Asn Pro Ala Asp  
 100 105 110



Phe Thr Val Leu Val Phe Phe Ala Pro Tyr Glu Lys Leu Asp Lys Arg  
115 120 125

Lys Lys Val Thr Lys Ala Leu Leu Gln Glu Ala Glu Ile Ile Asp Ala  
130 135 140

Ser Ser Pro Asp Gln Arg Asp Leu Lys Asp Met Val Gln Lys Lys Val  
145 150 155 160

Lys Ala Arg Gly Tyr Gln Phe Asp Lys Gly Ala Leu Lys Ala Leu Val  
165 170 175

Glu Lys Thr Asn Ala Asn Leu Ser Arg Val Met Gln Glu Leu Asp Lys  
180 185 190

Leu Phe Leu Tyr His Leu Asp Asp Lys Ile Ile Thr Val Gln Ser Val  
195 200 205

Asp Gln Val Val Ser Pro Ser Leu Glu Ser Asn Val Phe Ser Ile Asn  
210 215 220

Asp Tyr Ile Leu Ser Gly Gln Ser Gln Ala Ala Ile Arg Ala Phe Asn  
225 230 235 240

Asp Leu Ile Gln Gln Lys Glu Glu Pro Ile Lys Ile Ile Ala Ile Met  
245 250 255

Met Asn Gln Phe Arg Leu Leu Leu Gln Val Lys Ile Leu Arg Thr Lys  
260 265 270

Gly Tyr Gln Gln Gly Glu Ile Ala Lys Ile Leu Lys Val His Pro Tyr  
275 280 285

Arg Val Lys Leu Ala Ile Glu Lys Gln Glu Ile Phe Ser Lys Gln Ser  
290 295 300

Leu Ser Thr Ala Tyr Arg Tyr Leu Ile Glu Ser Asp His Leu Ile Lys  
305 310 315 320

Thr Gly Lys Val Thr Ser Gln Leu Gln Phe Glu Leu Phe Ala Leu Gln  
325 330 335

[illegible]

<210> 82  
<211> 158  
<212> PRT  
<213> Alloiococcus otitidis

<400> 82  
Met Asn Arg Ala Ile Tyr Ala Gly Ser Phe Asp Pro Ile Thr Leu Gly  
1 5 10 15

His Leu Asp Ile Ile Lys Arg Ala Ser His Leu Phe Asp Glu Val Ile  
20 25 30

Val Ala Val Ala Asn Asn Thr Ser Lys Asn Ser Met Leu Asn Phe Asp  
35 40 45

Gln Lys Leu Asn Leu Val Glu Gln Ser Ile Ala Ser Gln Gly Leu Ala  
50 55 60

Asn Val Gln Ala Lys Thr Leu Glu Ser Gly Leu Ile Val Asp Phe Ala  
65 70 75 80

Lys Asp Gln Gly Ala Ser Ser Leu Val Arg Gly Leu Arg Ser Val Lys  
85 90 95

Asp Phe Glu Tyr Glu Ile Ala Ile Glu Asp Leu Asn Lys Val Gln Asp  
100 105 110

Pro Ala Ile Glu Thr Val Tyr Leu Val Ser Ser Ser Lys Tyr Arg Ser  
115 120 125

Ile Ser Ser Ser Ile Val Arg Glu Ile Ile Lys Phe Asn Gly Arg Leu  
130 135 140

Asp Asp Leu Val Pro Asp Pro Val Val Glu Tyr Phe Lys Lys  
145 150 155

<210> 83  
<211> 1260  
<212> DNA  
<213> Alloiococcus otitidis

<220>  
<221> CDS  
<222> (28) .. (1260)  
<223>

<400> 83

ataaggattt caggaggaat actcata atg gat ttc aac tta gat aat aca gtt	54
Met Asp Phe Asn Leu Asp Asn Thr Val	
1 5	
tca ggt ggc gca aag att aag gtt att ggt gtt ggc ggt gct ggt ggc	102
Ser Gly Gly Ala Lys Ile Lys Val Ile Gly Val Gly Gly Ala Gly Gly	
10 15 20 25	
aat gcc gtt aac cgg atg att gaa gat gga gtc gaa ggc gtt gaa ttt	150
Asn Ala Val Asn Arg Met Ile Glu Asp Gly Val Glu Gly Val Glu Phe	
30 35 40	
att gta gcc aat aca gat gtc caa gcc ctt gat gcc aac cga gct gag	198
Ile Val Ala Asn Thr Asp Val Gln Ala Leu Asp Ala Asn Arg Ala Glu	
45 50 55	
act aaa att caa ctc gga gag aag tta acc agg gga ctc ggt gcc gga	246
Thr Lys Ile Gln Leu Gly Glu Lys Leu Thr Arg Gly Leu Gly Ala Gly	
60 65 70	
gct aat cca gaa gtt ggc cgt aag tcg gct gaa gag agt gaa gaa acc	294
Ala Asn Pro Glu Val Gly Arg Lys Ser Ala Glu Glu Ser Glu Glu Thr	
75 80 85	
att gcc gaa gct ctt gaa gga gct gac atg gtc ttc gtt act gct ggt	342
Ile Ala Glu Ala Leu Glu Gly Ala Asp Met Val Phe Val Thr Ala Gly	
90 95 100 105	
atg ggt ggc ggt act ggt act ggc ggg gcg ggc att att gcc cgc att	390
Met Gly Gly Gly Thr Gly Thr Gly Gly Ala Gly Ile Ile Ala Arg Ile	
110 115 120	
gcc aaa gaa caa ggg gct ttg act gta ggg gtt att acc cgg cgg ttc	438
Ala Lys Glu Gln Gly Ala Leu Thr Val Gly Val Ile Thr Arg Pro Phe	
125 130 135	
act ttt gaa gga cca aaa cgt ggg cgc ttt gca gcc gaa ggg att gcc	486
Thr Phe Glu Gly Pro Lys Arg Gly Arg Phe Ala Ala Glu Gly Ile Ala	
140 145 150	
caa atg cgg gaa cat gtt gac acc ctt gtc acc atc tcc aac aac cgc	534
Gln Met Arg Glu His Val Asp Thr Leu Val Thr Ile Ser Asn Asn Arg	
155 160 165	
ttg cta gaa att gtg gac aag aaa aca ccg atg atg gaa gcc ttc aga	582
Leu Leu Glu Ile Val Asp Lys Lys Thr Pro Met Met Glu Ala Phe Arg	
170 175 180 185	
gaa gca gat aat gtc ctc cgc caa ggg gtt caa ggt ata tct gac ttg	630
Glu Ala Asp Asn Val Leu Arg Gln Gly Val Gln Gly Ile Ser Asp Leu	
190 195 200	
att acc aat cca ggc tac gtc aac tta gac ttt gcc gat gtc aaa acg	678
Ile Thr Asn Pro Gly Tyr Val Asn Leu Asp Phe Ala Asp Val Lys Thr	
205 210 215	
gtg atg gcc aac caa ggt tct gcc ttg atg ggg att ggg tct gct tca	726

Val Met Ala Asn Gln Gly Ser Ala Leu Met Gly Ile Gly Ser Ala Ser  
 220 225 230

ggt gag aat aga acg gct gaa gct act aag aaa gct att tca tct cca 774  
 Gly Glu Asn Arg Thr Ala Glu Ala Thr Lys Lys Ala Ile Ser Ser Pro  
 235 240 245

ctt ttg gaa gtc tcc ctc aat ggg gct gaa aat gtc cta tta aac ata 822  
 Leu Leu Glu Val Ser Leu Asn Gly Ala Glu Asn Val Leu Leu Asn Ile  
 250 255 260 265

acc gga aac caa gac tta acc ctc ttt gaa gct caa gat gct tct gat 870  
 Thr Gly Asn Gln Asp Leu Thr Leu Phe Glu Ala Gln Asp Ala Ser Asp  
 270 275 280

atc gtc ggg gct gct gct tct ggt gat gtt aat att atc ttc ggt act 918  
 Ile Val Gly Ala Ala Ala Ser Gly Asp Val Asn Ile Ile Phe Gly Thr  
 285 290 295

tcc atc aat gaa gac ctg gaa gat gag gtc atc gtt acc gtt att gca 966  
 Ser Ile Asn Glu Asp Leu Glu Asp Glu Val Ile Val Thr Val Ile Ala  
 300 305 310

act ggt atc act ggt aaa gac atg ggc gag aaa tct tct aaa tcc tca 1014  
 Thr Gly Ile Thr Gly Lys Asp Met Gly Glu Lys Ser Ser Lys Ser Ser  
 315 320 325

aac cgt agc caa ggt cct agt caa aaa agt caa gct cga tca gct agt 1062  
 Asn Arg Ser Gln Gly Pro Ser Gln Lys Ser Gln Ala Arg Ser Ala Ser  
 330 335 340 345

gag tct agc ttc tct agc tgg caa aac caa tcc aat gaa aga cca ggg 1110  
 Glu Ser Ser Phe Ser Ser Trp Gln Asn Gln Ser Asn Glu Arg Pro Gly  
 350 355 360

gaa gac caa gac cga cca agc tct caa aga cgg gaa gtc gat cgg tcc 1158  
 Glu Asp Gln Asp Arg Pro Ser Ser Gln Arg Arg Glu Val Asp Arg Ser  
 365 370 375

gaa aac ctg ttc aat gac gat agt aag gac cag cca gca gac tct ggt 1206  
 Glu Asn Leu Phe Asn Asp Asp Ser Lys Asp Gln Pro Ala Asp Ser Gly  
 380 385 390

gat gat gac gaa ttg gat acc cct cct ttc ttt aga cgt cgc cgc aag 1254  
 Asp Asp Asp Glu Leu Asp Thr Pro Pro Phe Phe Arg Arg Arg Lys  
 395 400 405

aat tag 1260  
 Asn  
 410

&lt;210&gt; 84

&lt;211&gt; 410

&lt;212&gt; PRT

&lt;213&gt; Alloiococcus otitidis

&lt;400&gt; 84

Met Asp Phe Asn Leu Asp Asn Thr Val Ser Gly Gly Ala Lys Ile Lys  
1 5 10 15

Val Ile Gly Val Gly Gly Ala Gly Gly Asn Ala Val Asn Arg Met Ile  
20 25 30

Glu Asp Gly Val Glu Gly Val Glu Phe Ile Val Ala Asn Thr Asp Val  
35 40 45

Gln Ala Leu Asp Ala Asn Arg Ala Glu Thr Lys Ile Gln Leu Gly Glu  
50 55 60

Lys Leu Thr Arg Gly Leu Gly Ala Gly Ala Asn Pro Glu Val Gly Arg  
65 70 75 80

Lys Ser Ala Glu Glu Ser Glu Glu Thr Ile Ala Glu Ala Leu Glu Gly  
85 90 95

Ala Asp Met Val Phe Val Thr Ala Gly Met Gly Gly Gly Thr Gly Thr  
100 105 110

Gly Gly Ala Gly Ile Ile Ala Arg Ile Ala Lys Glu Gln Gly Ala Leu  
115 120 125

Thr Val Gly Val Ile Thr Arg Pro Phe Thr Phe Glu Gly Pro Lys Arg  
130 135 140

Gly Arg Phe Ala Ala Glu Gly Ile Ala Gln Met Arg Glu His Val Asp  
145 150 155 160

Thr Leu Val Thr Ile Ser Asn Asn Arg Leu Leu Glu Ile Val Asp Lys  
165 170 175

Lys Thr Pro Met Met Glu Ala Phe Arg Glu Ala Asp Asn Val Leu Arg  
180 185 190

Gln Gly Val Gln Gly Ile Ser Asp Leu Ile Thr Asn Pro Gly Tyr Val  
195 200 205

Asn Leu Asp Phe Ala Asp Val Lys Thr Val Met Ala Asn Gln Gly Ser  
210 215 220

Ala Leu Met Gly Ile Gly Ser Ala Ser Gly Glu Asn Arg Thr Ala Glu  
 225 230 235 240

Ala Thr Lys Lys Ala Ile Ser Ser Pro Leu Leu Glu Val Ser Leu Asn  
 245 250 255

Gly Ala Glu Asn Val Leu Leu Asn Ile Thr Gly Asn Gln Asp Leu Thr  
 260 265 270

Leu Phe Glu Ala Gln Asp Ala Ser Asp Ile Val Gly Ala Ala Ala Ser  
 275 280 285

Gly Asp Val Asn Ile Ile Phe Gly Thr Ser Ile Asn Glu Asp Leu Glu  
 290 295 300

Asp Glu Val Ile Val Thr Val Ile Ala Thr Gly Ile Thr Gly Lys Asp  
 305 310 315 320

Met Gly Glu Lys Ser Ser Lys Ser Ser Asn Arg Ser Gln Gly Pro Ser  
 325 330 335

Gln Lys Ser Gln Ala Arg Ser Ala Ser Glu Ser Ser Phe Ser Ser Trp  
 340 345 350

Gln Asn Gln Ser Asn Glu Arg Pro Gly Glu Asp Gln Asp Arg Pro Ser  
 355 360 365

Ser Gln Arg Arg Glu Val Asp Arg Ser Glu Asn Leu Phe Asn Asp Asp  
 370 375 380

Ser Lys Asp Gln Pro Ala Asp Ser Gly Asp Asp Asp Glu Leu Asp Thr  
 385 390 395 400

Pro Pro Phe Phe Arg Arg Arg Arg Lys Asn  
 405 410

<210> 85

<211> 1377

<212> DNA

<213> *Alloiococcus otitidis*

<220>

<221> CDS

<222> (13)..(1377)

<223>

<400> 85  
acagataata ga atg ttt tta gat atg gag gtt tca atg aat atg aaa aat 51  
Met Phe Leu Asp Met Glu Val Ser Met Asn Met Lys Asn  
1 5 10

ggg gtt tat aca agc ctt gat att gga acc act tca ata aaa gta gtt 99  
Gly Val Tyr Thr Ser Leu Asp Ile Gly Thr Thr Ser Ile Lys Val Val  
15 20 25

gtc agt gaa gtt gat aat aat cag ctc aaa gtt att gga gta gga aaa 147  
Val Ser Glu Val Asp Asn Asn Gln Leu Lys Val Ile Gly Val Gly Lys  
30 35 40 45

gct caa tca aaa ggt tta aaa agg ggc atg gtt gtc gat ata gat gct 195  
Ala Gln Ser Lys Gly Leu Lys Arg Gly Met Val Val Asp Ile Asp Ala  
50 55 60

acc gtc cag gcc att cat act gca gtg aag cag gct gct gat aag act 243  
Thr Val Gln Ala Ile His Thr Ala Val Lys Gln Ala Ala Asp Lys Thr  
65 70 75

ggg gtt atg atc aac cag ctc att gtt gga gtt cct gct aat ggt gtt 291  
Gly Val Met Ile Asn Gln Leu Ile Val Gly Val Pro Ala Asn Gly Val  
80 85 90

agt att gaa ccc tgt cac ggg gtc att act gta gat gac cgg tcc aag 339  
Ser Ile Glu Pro Cys His Gly Val Ile Thr Val Asp Asp Arg Ser Lys  
95 100 105

gaa ata gac agc cag gaa gtg aac cgg gta gtc aac cag tcc att gct 387  
Glu Ile Asp Ser Gln Glu Val Asn Arg Val Val Asn Gln Ser Ile Ala  
110 115 120 125

aat atc gtt ccg cca gat aga gac tta tta tcc gtc agt tta gaa gaa 435  
Asn Ile Val Pro Pro Asp Arg Asp Leu Leu Ser Val Ser Leu Glu Glu  
130 135 140

ttt att gta gat ggt ttt gat gaa att cat gat ccg aga ggc atg gtg 483  
Phe Ile Val Asp Gly Phe Asp Glu Ile His Asp Pro Arg Gly Met Val  
145 150 155

ggc cag cgg tta gaa ctt tac ggg aca gca att tca gtg cct aaa aca 531  
Gly Gln Arg Leu Glu Leu Tyr Gly Thr Ala Ile Ser Val Pro Lys Thr  
160 165 170

att tta cat aac att aga cgt tgt gtt gaa aaa gcg ggc tat caa att 579  
Ile Leu His Asn Ile Arg Arg Cys Val Glu Lys Ala Gly Tyr Gln Ile  
175 180 185

gct gcc tta att ctc cag ccc caa gcc atg gcc aag gta gcc ttg tct 627  
Ala Ala Leu Ile Leu Gln Pro Gln Ala Met Ala Lys Val Ala Leu Ser  
190 195 200 205

gag gat gag cgg aat ttt ggt aca gtt atg gtg gat ata ggc gga ggt 675  
Glu Asp Glu Arg Asn Phe Gly Thr Val Met Val Asp Ile Gly Gly Gly  
210 215 220



caa acg acc cta tca gcc att cac gat gag caa gtg aag tat gcc aat Gln Thr Thr Leu Ser Ala Ile His Asp Glu Gln Val Lys Tyr Ala Asn 225 230 235	723
gtg gtc caa gaa gcc gga gaa tat att acc aaa gac att tcc att gtc Val Val Gln Glu Ala Gly Glu Tyr Ile Thr Lys Asp Ile Ser Ile Val 240 245 250	771
atc aac acc tca cag caa aat gca gaa aag ctc aaa aga gaa gtt ggg Ile Asn Thr Ser Gln Gln Asn Ala Glu Lys Leu Lys Arg Glu Val Gly 255 260 265	819
gcc att aaa agt cag tct gat tca act gtt caa gta gat gtt gta ggt Ala Ile Lys Ser Gln Ser Asp Ser Thr Val Gln Val Asp Val Val Gly 270 275 280 285	867
caa aat gaa cct gtg aag att aaa gaa tcc tat gtc ggt gaa att att Gln Asn Glu Pro Val Lys Ile Lys Glu Ser Tyr Val Gly Glu Ile Ile 290 295 300	915
gaa gcc cgg gtt agc caa atc ttt gaa aaa gtg aag gct gac ctt gac Glu Ala Arg Val Ser Gln Ile Phe Glu Lys Val Lys Ala Asp Leu Asp 305 310 315	963
cca att aac gcc ttc caa ttg cca ggt ggt gcc gtt att tcc ggc ggt Pro Ile Asn Ala Phe Gln Leu Pro Gly Gly Ala Val Ile Ser Gly Gly 320 325 330	1011
tca gct gcc ata cca ggt att gac agc ttg gct gaa gac atc ttc aag Ser Ala Ala Ile Pro Gly Ile Asp Ser Leu Ala Glu Asp Ile Phe Lys 335 340 345	1059
gtt cgg tca gag ctc tac att ccc gac tac atg ggt atc cga act ccc Val Arg Ser Glu Leu Tyr Ile Pro Asp Tyr Met Gly Ile Arg Thr Pro 350 355 360 365	1107
gcc ttc act gtg gca gtc ggc ttg acc ctc tac caa gcc cag act tct Ala Phe Thr Val Ala Val Gly Leu Thr Leu Tyr Gln Ala Gln Thr Ser 370 375 380	1155
gat att gag cgg gcc atc aac cag tcc atc ttg caa aat atc ggt att Asp Ile Glu Arg Ala Ile Asn Gln Ser Ile Leu Gln Asn Ile Gly Ile 385 390 395	1203
aat cca gat agc cag cct gct aac cgg ata gtt gac cag gat gat tca Asn Pro Asp Ser Gln Pro Ala Asn Arg Ile Val Asp Gln Asp Asp Ser 400 405 410	1251
gtc caa agt cag gac caa aag acg caa gat gag cca gca gga gac caa Val Gln Ser Gln Asp Gln Lys Thr Gln Asp Glu Pro Ala Gly Asp Gln 415 420 425	1299
gct agt cag tcg gat agt cca gaa gaa ggc aat ttt aca gac aga atc Ala Ser Gln Ser Asp Ser Pro Glu Glu Gly Asn Phe Thr Asp Arg Ile 430 435 440 445	1347

aag cat ttc ttt act aca ttt ttc gat taa  
 Lys His Phe Phe Thr Thr Phe Phe Asp  
 450

1377

<210> 86  
 <211> 454  
 <212> PRT  
 <213> Alloiococcus otitidis

<400> 86  
 Met Phe Leu Asp Met Glu Val Ser Met Asn Met Lys Asn Gly Val Tyr  
 1 5 10 15

Thr Ser Leu Asp Ile Gly Thr Thr Ser Ile Lys Val Val Val Ser Glu  
 20 25 30

Val Asp Asn Asn Gln Leu Lys Val Ile Gly Val Gly Lys Ala Gln Ser  
 35 40 45

Lys Gly Leu Lys Arg Gly Met Val Val Asp Ile Asp Ala Thr Val Gln  
 50 55 60

Ala Ile His Thr Ala Val Lys Gln Ala Ala Asp Lys Thr Gly Val Met  
 65 70 75 80

Ile Asn Gln Leu Ile Val Gly Val Pro Ala Asn Gly Val Ser Ile Glu  
 85 90 95

Pro Cys His Gly Val Ile Thr Val Asp Asp Arg Ser Lys Glu Ile Asp  
 100 105 110

Ser Gln Glu Val Asn Arg Val Val Asn Gln Ser Ile Ala Asn Ile Val  
 115 120 125

Pro Pro Asp Arg Asp Leu Leu Ser Val Ser Leu Glu Glu Phe Ile Val  
 130 135 140

Asp Gly Phe Asp Glu Ile His Asp Pro Arg Gly Met Val Gly Gln Arg  
 145 150 155 160

Leu Glu Leu Tyr Gly Thr Ala Ile Ser Val Pro Lys Thr Ile Leu His  
 165 170 175

Asn Ile Arg Arg Cys Val Glu Lys Ala Gly Tyr Gln Ile Ala Ala Leu  
 180 185 190

Ile Leu Gln Pro Gln Ala Met Ala Lys Val Ala Leu Ser Glu Asp Glu  
195 200 205

Arg Asn Phe Gly Thr Val Met Val Asp Ile Gly Gly Gly Gln Thr Thr  
210 215 220

Leu Ser Ala Ile His Asp Glu Gln Val Lys Tyr Ala Asn Val Val Gln  
225 230 235 240

Glu Ala Gly Glu Tyr Ile Thr Lys Asp Ile Ser Ile Val Ile Asn Thr  
245 250 255

Ser Gln Gln Asn Ala Glu Lys Leu Lys Arg Glu Val Gly Ala Ile Lys  
260 265 270

Ser Gln Ser Asp Ser Thr Val Gln Val Asp Val Val Gly Gln Asn Glu  
275 280 285

Pro Val Lys Ile Lys Glu Ser Tyr Val Gly Glu Ile Ile Glu Ala Arg  
290 295 300

Val Ser Gln Ile Phe Glu Lys Val Lys Ala Asp Leu Asp Pro Ile Asn  
305 310 315 320

Ala Phe Gln Leu Pro Gly Gly Ala Val Ile Ser Gly Gly Ser Ala Ala  
325 330 335

Ile Pro Gly Ile Asp Ser Leu Ala Glu Asp Ile Phe Lys Val Arg Ser  
340 345 350

Glu Leu Tyr Ile Pro Asp Tyr Met Gly Ile Arg Thr Pro Ala Phe Thr  
355 360 365

Val Ala Val Gly Leu Thr Leu Tyr Gln Ala Gln Thr Ser Asp Ile Glu  
370 375 380

Arg Ala Ile Asn Gln Ser Ile Leu Gln Asn Ile Gly Ile Asn Pro Asp  
385 390 395 400

Ser Gln Pro Ala Asn Arg Ile Val Asp Gln Asp Asp Ser Val Gln Ser  
405 410 415

Gln Asp Gln Lys Thr Gln Asp Glu Pro Ala Gly Asp Gln Ala Ser Gln  
 420 425 430

Ser Asp Ser Pro Glu Glu Gly Asn Phe Thr Asp Arg Ile Lys His Phe  
 435 440 445

Phe Thr Thr Phe Phe Asp  
 450

<210> 87

<211> 1179

<212> DNA

<213> *Alloiococcus otitidis*

<220>

<221> CDS

<222> (16) .. (1179)

<223>

<400> 87

agcaaaggag caagt atg gaa act aaa aaa caa gca tta aaa gtt tta tta 51  
 Met Glu Thr Lys Lys Gln Ala Leu Lys Val Leu Leu  
 1 5 10

tca ggc ggt gga aca ggt ggc cat atc tac cca gcc ttg gcc ctt gct 99  
 Ser Gly Gly Gly Thr Gly Gly His Ile Tyr Pro Ala Leu Ala Leu Ala  
 15 20 25

aag cac cta gct agc tta cac tca gat gtc gag ttt ttg tat gtt ggc 147  
 Lys His Leu Ala Ser Leu His Ser Asp Val Glu Phe Leu Tyr Val Gly  
 30 35 40

act caa agg gga ttg gaa aat aaa ttg gtc ccc caa gca gga ctt gac 195  
 Thr Gln Arg Gly Leu Glu Asn Lys Leu Val Pro Gln Ala Gly Leu Asp  
 45 50 55 60

ttt atc ccg atc aaa gta gaa gga ttt agc cgg aag ttt aac ttc aaa 243  
 Phe Ile Pro Ile Lys Val Glu Gly Phe Ser Arg Lys Phe Asn Phe Lys  
 65 70 75

agc att aaa tat aat act aaa agt ctg att tat ttt cta aag gcc ctg 291  
 Ser Ile Lys Tyr Asn Thr Lys Ser Leu Ile Tyr Phe Leu Lys Ala Leu  
 80 85 90

agt aag tct aag caa atc atc aaa gac ttt cag cca gat gtg gta ata 339  
 Ser Lys Ser Lys Gln Ile Ile Lys Asp Phe Gln Pro Asp Val Val Ile  
 95 100 105

ggg aca ggt ggt tat gtt tgt gcc cct gtc ata tac cag gcg acc aag 387  
 Gly Thr Gly Gly Tyr Val Cys Ala Pro Val Ile Tyr Gln Ala Thr Lys  
 110 115 120

tta ggc att cca agt ctc att cac gaa caa aat agt gtc gcc ggg gtg 435

Leu Gly Ile Pro Ser	Leu Ile His Glu Gln Asn Ser Val Ala Gly Val	
125	130 135 140	
acc aat aag ttt ttg gct cgg tac gta gac aag att gcc cta agt ttc		483
Thr Asn Lys Phe Leu Ala Arg Tyr Val Asp Lys Ile Ala Leu Ser Phe	145 150 155	
cag gaa gct gaa aaa tcc ttt gcc aag tat aag gat aag ctg gtt ttg		531
Gln Glu Ala Glu Lys Ser Phe Ala Lys Tyr Lys Asp Lys Leu Val Leu	160 165 170	
act ggt aat cca aga gga cag gaa gtc agc caa gtc aag ggt ggc ctt		579
Thr Gly Asn Pro Arg Gly Gln Glu Val Ser Gln Val Lys Gly Gly Leu	175 180 185	
agc ctc cac aag tat ggc atg gac atg tcc caa cct tca gta att att		627
Ser Leu His Lys Tyr Gly Met Asp Met Ser Gln Pro Ser Val Ile Ile	190 195 200	
ttt ggt ggg tca agg ggg gct tat gct att aat aag gcc ttt gtt gag		675
Phe Gly Gly Ser Arg Gly Ala Tyr Ala Ile Asn Lys Ala Phe Val Glu	205 210 215 220	
gca tat agt caa ctg gct gag agg gac tac cag gtc ttg ttt gtg ccg		723
Ala Tyr Ser Gln Leu Ala Glu Arg Asp Tyr Gln Val Leu Phe Val Pro	225 230 235	
gga tca gct aat ttt agc cgg ata aaa cag gaa att gat aac cgc tat		771
Gly Ser Ala Asn Phe Ser Arg Ile Lys Gln Glu Ile Asp Asn Arg Tyr	240 245 250	
ggc cag cat aag ccg tca aac att ttt att gaa tcc tat atc gat aac		819
Gly Gln His Lys Pro Ser Asn Ile Phe Ile Glu Ser Tyr Ile Asp Asn	255 260 265	
atg ccc caa gtt ttt aag gct att gac ttg gtg gtt tgc cgt agt ggg		867
Met Pro Gln Val Phe Lys Ala Ile Asp Leu Val Val Cys Arg Ser Gly	270 275 280	
gcc act acc cta gcc gaa att atg tca tta ggc ttg gcc agc att tta		915
Ala Thr Thr Leu Ala Glu Ile Met Ser Leu Gly Leu Ala Ser Ile Leu	285 290 295 300	
att cca agt ccc aat gta acg gct gac cac caa acc aaa aat gct atg		963
Ile Pro Ser Pro Asn Val Thr Ala Asp His Gln Thr Lys Asn Ala Met	305 310 315	
agt ttg gtt aac caa caa gct ggc tta atg att aag gaa aat gat cta		1011
Ser Leu Val Asn Gln Gln Ala Gly Leu Met Ile Lys Glu Asn Asp Leu	320 325 330	
aat ggc caa agc ctc tta aac tgc tta gat gac ctg atg cat gat gac		1059
Asn Gly Gln Ser Leu Leu Asn Cys Leu Asp Asp Leu Met His Asp Asp	335 340 345	
gca aaa aga aac aag atg gcc caa caa gcg aaa gaa atg ggc caa ccc		1107
Ala Lys Arg Asn Lys Met Ala Gln Gln Ala Lys Glu Met Gly Gln Pro		

350

355

360

caa gct tca gac aag ttg atc gct ctc atc ttg tcc atg gtt aag gaa  
 Gln Ala Ser Asp Lys Leu Ile Ala Leu Ile Leu Ser Met Val Lys Glu  
 365 370 375 380

1155

gat att aac tca gac atc gat taa  
 Asp Ile Asn Ser Asp Ile Asp  
 385

1179

&lt;210&gt; 88

&lt;211&gt; 387

&lt;212&gt; PRT

&lt;213&gt; Alloioococcus otitidis

&lt;400&gt; 88

Met Glu Thr Lys Lys Gln Ala Leu Lys Val Leu Leu Ser Gly Gly Gly  
 1 5 10 15

Thr Gly Gly His Ile Tyr Pro Ala Leu Ala Leu Ala Lys His Leu Ala  
 20 25 30

Ser Leu His Ser Asp Val Glu Phe Leu Tyr Val Gly Thr Gln Arg Gly  
 35 40 45

Leu Glu Asn Lys Leu Val Pro Gln Ala Gly Leu Asp Phe Ile Pro Ile  
 50 55 60

Lys Val Glu Gly Phe Ser Arg Lys Phe Asn Phe Lys Ser Ile Lys Tyr  
 65 70 75 80

Asn Thr Lys Ser Leu Ile Tyr Phe Leu Lys Ala Leu Ser Lys Ser Lys  
 85 90 95

Gln Ile Ile Lys Asp Phe Gln Pro Asp Val Val Ile Gly Thr Gly Gly  
 100 105 110

Tyr Val Cys Ala Pro Val Ile Tyr Gln Ala Thr Lys Leu Gly Ile Pro  
 115 120 125

Ser Leu Ile His Glu Gln Asn Ser Val Ala Gly Val Thr Asn Lys Phe  
 130 135 140

Leu Ala Arg Tyr Val Asp Lys Ile Ala Leu Ser Phe Gln Glu Ala Glu  
 145 150 155 160

Lys Ser Phe Ala Lys Tyr Lys Asp Lys Leu Val Leu Thr Gly Asn Pro  
165 170 175

Arg Gly Gln Glu Val Ser Gln Val Lys Gly Gly Leu Ser Leu His Lys  
180 185 190

Tyr Gly Met Asp Met Ser Gln Pro Ser Val Ile Ile Phe Gly Gly Ser  
195 200 205

Arg Gly Ala Tyr Ala Ile Asn Lys Ala Phe Val Glu Ala Tyr Ser Gln  
210 215 220

Leu Ala Glu Arg Asp Tyr Gln Val Leu Phe Val Pro Gly Ser Ala Asn  
225 230 235 240

Phe Ser Arg Ile Lys Gln Glu Ile Asp Asn Arg Tyr Gly Gln His Lys  
245 250 255

Pro Ser Asn Ile Phe Ile Glu Ser Tyr Ile Asp Asn Met Pro Gln Val  
260 265 270

Phe Lys Ala Ile Asp Leu Val Val Cys Arg Ser Gly Ala Thr Thr Leu  
275 280 285

Ala Glu Ile Met Ser Leu Gly Leu Ala Ser Ile Leu Ile Pro Ser Pro  
290 295 300

Asn Val Thr Ala Asp His Gln Thr Lys Asn Ala Met Ser Leu Val Asn  
305 310 315 320

Gln Gln Ala Gly Leu Met Ile Lys Glu Asn Asp Leu Asn Gly Gln Ser  
325 330 335

Leu Leu Asn Cys Leu Asp Asp Leu Met His Asp Asp Ala Lys Arg Asn  
340 345 350

Lys Met Ala Gln Gln Ala Lys Glu Met Gly Gln Pro Gln Ala Ser Asp  
355 360 365

Lys Leu Ile Ala Leu Ile Leu Ser Met Val Lys Glu Asp Ile Asn Ser  
370 375 380

Asp Ile Asp





gtt gtc gaa ctc tct agt ttc caa cta cag tca ctg acc tat ttc acc Val Val Glu Leu Ser Ser Phe Gln Leu Gln Ser Leu Thr Tyr Phe Thr 170 175 180 185	579
cct cat ata gca gtc att acc aat att tat tcc gcc cac ctt gac tac Pro His Ile Ala Val Ile Thr Asn Ile Tyr Ser Ala His Leu Asp Tyr 190 195 200	627
cat aag agt cgg gag gaa tat gtt agg gct aag cta agg att acc cag His Lys Ser Arg Glu Glu Tyr Val Arg Ala Lys Leu Arg Ile Thr Gln 205 210 215	675
gct caa ggt cgg gat gac tac cta gtc tac tac cag ggt cag gaa gaa Ala Gln Gly Pro Asp Asp Tyr Leu Val Tyr Tyr Gln Gly Gln Glu Glu 220 225 230	723
ttg gct agc ctg gtc aaa aaa tac tct aaa gcc cag ctg gtc ccc tat Leu Ala Ser Leu Val Lys Lys Tyr Ser Lys Ala Gln Leu Val Pro Tyr 235 240 245	771
act gac aag ggt caa ctg aac caa gga gcc tat atc aag gat gac tat Thr Asp Lys Gly Gln Leu Asn Gln Gly Ala Tyr Ile Lys Asp Asp Tyr 250 255 260 265	819
ctt atc tat aat caa gag cca gtc atg gct tta gac cga gtt caa gtt Leu Ile Tyr Asn Gln Glu Pro Val Met Ala Leu Asp Arg Val Gln Val 270 275 280	867
tct ggt agc cac aac tta caa aat att tta gca gct gtt tgc gta gct Ser Gly Ser His Asn Leu Gln Asn Ile Leu Ala Ala Val Cys Val Ala 285 290 295	915
aaa ata aag ggg ctc tct aac caa acc att gcc caa gct gtc aac cac Lys Ile Lys Gly Leu Ser Asn Gln Thr Ile Ala Gln Ala Val Asn His 300 305 310	963
ttc aaa ggg gtt gcc cac cgc agc cag gtg gtt ggg cgg tat gag gac Phe Lys Gly Val Ala His Arg Ser Gln Val Val Gly Arg Tyr Glu Asp 315 320 325	1011
cgg ctt ttt gtc aac gac tct aag gca acc aat agc ttg gcc aca cag Arg Leu Phe Val Asn Asp Ser Lys Ala Thr Asn Ser Leu Ala Thr Gln 330 335 340 345	1059
aag gca tta gaa gcc tat gac caa gat acc atc ttg tta gtg ggt ggc Lys Ala Leu Glu Ala Tyr Asp Gln Asp Thr Ile Leu Leu Val Gly Gly 350 355 360	1107
cta gac cgc caa gat gat ttt tcc aag ctt gac cat gct cta aac agg Leu Asp Arg Gln Asp Asp Phe Ser Lys Leu Asp His Ala Leu Asn Arg 365 370 375	1155
gtt aag ggg gtc gtt tgt ttt ggc cag acc aaa gat aag tta gcc cgg Val Lys Gly Val Val Cys Phe Gly Gln Thr Lys Asp Lys Leu Ala Arg 380 385 390	1203

tat ttt aaa gac cgt cac att gag ggt gtt gag ctt gcc cag aca gtt 1251  
 Tyr Phe Lys Asp Arg His Ile Glu Gly Val Glu Leu Ala Gln Thr Val  
 395 400 405

cct gaa gca gtt gat ttg gct tac gac ttg agt gag cca gga caa gtc 1299  
 Pro Glu Ala Val Asp Leu Ala Tyr Asp Leu Ser Glu Pro Gly Gln Val  
 410 415 420 425

att tta ttt tct cct gct tgt gca agt tgg gac caa tat gct aac ttt 1347  
 Ile Leu Phe Ser Pro Ala Cys Ala Ser Trp Asp Gln Tyr Ala Asn Phe  
 430 435 440

gaa gag aga gga caa gat tat gtt gat gca atc cag cag ctg gtt gaa 1395  
 Glu Glu Arg Gly Gln Asp Tyr Val Asp Ala Ile Gln Gln Leu Val Glu  
 445 450 455

aga cta gag caa agg agc aag tat gga aac taa 1428  
 Arg Leu Glu Gln Arg Ser Lys Tyr Gly Asn  
 460 465

<210> 90  
 <211> 467  
 <212> PRT  
 <213> *Alloiococcus otitidis*

<400> 90  
 Met Val Asp Ser Val Phe Cys Asn Lys Lys Val Leu Val Leu Gly Leu  
 1 5 10 15

Ala Lys Ser Gly Leu Ser Ala Ala His Leu Leu Lys Lys Leu Gly Ala  
 20 25 30

Lys Val Ile Val Asn Asp Lys Leu Ala Leu Glu Asn Asn Thr Glu Ala  
 35 40 45

Gln Val Leu Ile Glu Glu Gly Phe Gln Val Ile Thr Gly Tyr His Pro  
 50 55 60

Glu Asp Leu Leu Asp Ala Ser Phe Asp Phe Val Val Lys Asn Pro Gly  
 65 70 75 80

Ile Pro Tyr Thr Asn Pro Val Val Gly Gln Ala Glu Lys Leu Ala Ile  
 85 90 95

Pro Ile Leu Thr Glu Val Asp Val Ala Gly Ser Ile Leu Lys Ala Lys  
 100 105 110

Pro Ile Ala Val Thr Gly Thr Asn Gly Lys Thr Thr Thr Val Ser Leu  
 115 120 125

Ile Tyr Asp Ile Leu Ala Gln Asp Gln Ala Glu Ser Pro Glu Pro Lys  
130 135 140

Pro Val Tyr Lys Leu Gly Asn Ile Gly Gln Pro Val Ser Asp Leu Ala  
145 150 155 160

Leu Glu Ile Lys Ala Glu Ser Asn Leu Val Val Glu Leu Ser Ser Phe  
165 170 175

Gln Leu Gln Ser Leu Thr Tyr Phe Thr Pro His Ile Ala Val Ile Thr  
180 185 190

Asn Ile Tyr Ser Ala His Leu Asp Tyr His Lys Ser Arg Glu Glu Tyr  
195 200 205

Val Arg Ala Lys Leu Arg Ile Thr Gln Ala Gln Gly Pro Asp Asp Tyr  
210 215 220

Leu Val Tyr Tyr Gln Gly Gln Glu Glu Leu Ala Ser Leu Val Lys Lys  
225 230 235 240

Tyr Ser Lys Ala Gln Leu Val Pro Tyr Thr Asp Lys Gly Gln Leu Asn  
245 250 255

Gln Gly Ala Tyr Ile Lys Asp Asp Tyr Leu Ile Tyr Asn Gln Glu Pro  
260 265 270

Val Met Ala Leu Asp Arg Val Gln Val Ser Gly Ser His Asn Leu Gln  
275 280 285

Asn Ile Leu Ala Ala Val Cys Val Ala Lys Ile Lys Gly Leu Ser Asn  
290 295 300

Gln Thr Ile Ala Gln Ala Val Asn His Phe Lys Gly Val Ala His Arg  
305 310 315 320

Ser Gln Val Val Gly Arg Tyr Glu Asp Arg Leu Phe Val Asn Asp Ser  
325 330 335

Lys Ala Thr Asn Ser Leu Ala Thr Gln Lys Ala Leu Glu Ala Tyr Asp  
340 345 350

Gln Asp Thr Ile Leu Leu Val Gly Gly Leu Asp Arg Gln Asp Asp Phe  
 355 360 365

Ser Lys Leu Asp His Ala Leu Asn Arg Val Lys Gly Val Val Cys Phe  
 370 375 380

Gly Gln Thr Lys Asp Lys Leu Ala Arg Tyr Phe Lys Asp Arg His Ile  
 385 390 395 400

Glu Gly Val Glu Leu Ala Gln Thr Val Pro Glu Ala Val Asp Leu Ala  
 405 410 415

Tyr Asp Leu Ser Glu Pro Gly Gln Val Ile Leu Phe Ser Pro Ala Cys  
 420 425 430

Ala Ser Trp Asp Gln Tyr Ala Asn Phe Glu Glu Arg Gly Gln Asp Tyr  
 435 440 445

Val Asp Ala Ile Gln Gln Leu Val Glu Arg Leu Glu Gln Arg Ser Lys  
 450 455 460

Tyr Gly Asn  
 465

<210> 91  
 <211> 651  
 <212> DNA  
 <213> *Alloiococcus otitidis*

<220>  
 <221> CDS  
 <222> (7)..(651)  
 <223>

<400> 91  
 actagt atg aag caa aaa act caa gcg aca gcg gtc aac cag acc caa 48  
 Met Lys Gln Lys Thr Gln Ala Thr Ala Val Asn Gln Thr Gln  
 1 5 10

aca gag gca gaa gaa aga caa gaa acc cgt cgg aaa att ggc ctc atg 96  
 Thr Glu Ala Glu Glu Arg Gln Glu Thr Arg Arg Lys Ile Gly Leu Met  
 15 20 25 30

ggg ggg acc ttt aat ccg ccc cat ctg ggt cat tta ctg gta gct gaa 144  
 Gly Gly Thr Phe Asn Pro Pro His Leu Gly His Leu Leu Val Ala Glu  
 35 40 45

caa gtt tat gag gcc ttg gac ttg gat aat att cac ttt atg ccc act 192

Gln Val Tyr Glu Ala Leu Asp Leu Asp Asn Ile His Phe Met Pro Thr  
 50 55 60  
 gca aag ccg ggc cat gcc gct ggt aag gaa acc ata gat gcc tct tac 240  
 Ala Lys Pro Gly His Ala Ala Gly Lys Glu Thr Ile Asp Ala Ser Tyr  
 65 70 75  
 cgg gtt gat atg gtg gat tat gcc atc gaa gat aac ccc cac ttt tct 288  
 Arg Val Asp Met Val Asp Tyr Ala Ile Glu Asp Asn Pro His Phe Ser  
 80 85 90  
 ctt aac ttg act gaa gtg aac cgg gga ggg aca act tac acc atc gat 336  
 Leu Asn Leu Thr Glu Val Asn Arg Gly Gly Thr Thr Tyr Thr Ile Asp  
 95 100 105 110  
 acc att aaa gaa ttg aaa gag gct agc ccg aat aca gat tat tac ttc 384  
 Thr Ile Lys Glu Leu Lys Glu Ala Ser Pro Asn Thr Asp Tyr Tyr Phe  
 115 120 125  
 att att ggt gag gat tca gtt atg gat ttg gcc cag tgg aag aat att 432  
 Ile Ile Gly Glu Asp Ser Val Met Asp Leu Ala Gln Trp Lys Asn Ile  
 130 135 140  
 gaa caa tta ctg gat tta gtt caa ttt gtt ggt gtg aag cga cca ggc 480  
 Glu Gln Leu Leu Asp Leu Val Gln Phe Val Gly Val Lys Arg Pro Gly  
 145 150 155  
 tac caa gct gat gtg gac ttt ccc att att tgg gtg gat acg cca gaa 528  
 Tyr Gln Ala Asp Val Asp Phe Pro Ile Ile Trp Val Asp Thr Pro Glu  
 160 165 170  
 cta gat att agt tca agt gac atc agg caa agg gtg gca gaa ggg caa 576  
 Leu Asp Ile Ser Ser Ser Asp Ile Arg Gln Arg Val Ala Glu Gly Gln  
 175 180 185 190  
 tcc att aaa tat ttg acc cca gat agg gta aga gat tat att gaa gac 624  
 Ser Ile Lys Tyr Leu Thr Pro Asp Arg Val Arg Asp Tyr Ile Glu Asp  
 195 200 205  
 aat ggc tta tat aag ggt gaa gaa taa 651  
 Asn Gly Leu Tyr Lys Gly Glu Glu  
 210

&lt;210&gt; 92

&lt;211&gt; 214

&lt;212&gt; PRT

<213> *Alloiococcus otitidis*

&lt;400&gt; 92

Met Lys Gln Lys Thr Gln Ala Thr Ala Val Asn Gln Thr Gln Thr Glu  
 1 5 10 15

Ala Glu Glu Arg Gln Glu Thr Arg Arg Lys Ile Gly Leu Met Gly Gly  
 20 25 30

Thr Phe Asn Pro Pro His Leu Gly His Leu Leu Val Ala Glu Gln Val  
 35 40 45  
 Tyr Glu Ala Leu Asp Leu Asp Asn Ile His Phe Met Pro Thr Ala Lys  
 50 55 60  
 Pro Gly His Ala Ala Gly Lys Glu Thr Ile Asp Ala Ser Tyr Arg Val  
 65 70 75 80  
 Asp Met Val Asp Tyr Ala Ile Glu Asp Asn Pro His Phe Ser Leu Asn  
 85 90 95  
 Leu Thr Glu Val Asn Arg Gly Gly Thr Thr Tyr Thr Ile Asp Thr Ile  
 100 105 110  
 Lys Glu Leu Lys Glu Ala Ser Pro Asn Thr Asp Tyr Tyr Phe Ile Ile  
 115 120 125  
 Gly Glu Asp Ser Val Met Asp Leu Ala Gln Trp Lys Asn Ile Glu Gln  
 130 135 140  
 Leu Leu Asp Leu Val Gln Phe Val Gly Val Lys Arg Pro Gly Tyr Gln  
 145 150 155 160  
 Ala Asp Val Asp Phe Pro Ile Ile Trp Val Asp Thr Pro Glu Leu Asp  
 165 170 175  
 Ile Ser Ser Ser Asp Ile Arg Gln Arg Val Ala Glu Gly Gln Ser Ile  
 180 185 190  
 Lys Tyr Leu Thr Pro Asp Arg Val Arg Asp Tyr Ile Glu Asp Asn Gly  
 195 200 205  
 Leu Tyr Lys Gly Glu Glu  
 210

&lt;210&gt; 93

&lt;211&gt; 666

&lt;212&gt; DNA

<213> *Alloiococcus otitidis*

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(666)

&lt;223&gt;

&lt;400&gt; 93

atg gta ggg gga ctc att ttt gtc ctc act gcc agc aat aaa agg aaa	48
Met Val Gly Gly Leu Ile Phe Val Leu Thr Ala Ser Asn Lys Arg Lys	
1 5 10 15	
gga agt ttg tcc atg acc tat ttg tta ggc cta acc ggt ggc att gcc	96
Gly Ser Leu Ser Met Thr Tyr Leu Leu Gly Leu Thr Gly Gly Ile Ala	
20 25 30	
agt ggg aag tct act gtt agc cag gtt ttt aag gaa aag ggt atc caa	144
Ser Gly Lys Ser Thr Val Ser Gln Val Phe Lys Glu Lys Gly Ile Gln	
35 40 45	
gtg gtt gat gct gac cga gtt gcc cga cag gtt gtt gaa cct gga agt	192
Val Val Asp Ala Asp Arg Val Ala Arg Gln Val Val Glu Pro Gly Ser	
50 55 60	
cca ggc tta gac cag ctt gtt gat tat ttt ggc cag gag att ttg acc	240
Pro Gly Leu Asp Gln Leu Val Asp Tyr Phe Gly Gln Glu Ile Leu Thr	
65 70 75 80	
cag gat ggg ggc ttg gac cgc aaa tat tta ggc gac ctt atc ttc cgg	288
Gln Asp Gly Gly Leu Asp Arg Lys Tyr Leu Gly Asp Leu Ile Phe Arg	
85 90 95	
aat agc cag gcc aag gag gct gtc aac cgg atc ctc cac cct ttg att	336
Asn Ser Gln Ala Lys Glu Ala Val Asn Arg Ile Leu His Pro Leu Ile	
100 105 110	
agg cag tct atc caa aat caa att aaa act gcc ata ggc caa gac ttg	384
Arg Gln Ser Ile Gln Asn Gln Ile Lys Thr Ala Ile Gly Gln Asp Leu	
115 120 125	
gat ttg tta gtt tta gac atc ccc ctc ctt tac gag aca ggt cag gca	432
Asp Leu Leu Val Leu Asp Ile Pro Leu Leu Tyr Glu Thr Gly Gln Ala	
130 135 140	
gac gac tac cag gcc gtc atg gtg gtt tcg ctt ccc tac cag gac cag	480
Asp Asp Tyr Gln Ala Val Met Val Val Ser Leu Pro Tyr Gln Asp Gln	
145 150 155 160	
gtg agt cgg tta atg gac cgg gat ggg att gac cga gac caa gcc ctg	528
Val Ser Arg Leu Met Asp Arg Asp Gly Ile Asp Arg Asp Gln Ala Leu	
165 170 175	
cgc aag att cag gcc caa atg tca ttg gaa gaa aaa gtg aag ttg gcg	576
Arg Lys Ile Gln Ala Gln Met Ser Leu Glu Glu Lys Val Lys Leu Ala	
180 185 190	
gac tat gtc att gat aac agc gga agc aag gaa gaa agc cgt cag cag	624
Asp Tyr Val Ile Asp Asn Ser Gly Ser Lys Glu Glu Ser Arg Gln Gln	
195 200 205	
gtt gaa gct tgg ttg gat caa aag ggt ttt aaa aac ttg taa	666
Val Glu Ala Trp Leu Asp Gln Lys Gly Phe Lys Asn Leu	

210

215

220

&lt;210&gt; 94

&lt;211&gt; 221

&lt;212&gt; PRT

<213> *Alloiococcus otitidis*

&lt;400&gt; 94

Met Val Gly Gly Leu Ile Phe Val Leu Thr Ala Ser Asn Lys Arg Lys  
1 5 10 15

Gly Ser Leu Ser Met Thr Tyr Leu Leu Gly Leu Thr Gly Gly Ile Ala  
20 25 30

Ser Gly Lys Ser Thr Val Ser Gln Val Phe Lys Glu Lys Gly Ile Gln  
35 40 45

Val Val Asp Ala Asp Arg Val Ala Arg Gln Val Val Glu Pro Gly Ser  
50 55 60

Pro Gly Leu Asp Gln Leu Val Asp Tyr Phe Gly Gln Glu Ile Leu Thr  
65 70 75 80

Gln Asp Gly Gly Leu Asp Arg Lys Tyr Leu Gly Asp Leu Ile Phe Arg  
85 90 95

Asn Ser Gln Ala Lys Glu Ala Val Asn Arg Ile Leu His Pro Leu Ile  
100 105 110

Arg Gln Ser Ile Gln Asn Gln Ile Lys Thr Ala Ile Gly Gln Asp Leu  
115 120 125

Asp Leu Leu Val Leu Asp Ile Pro Leu Leu Tyr Glu Thr Gly Gln Ala  
130 135 140

Asp Asp Tyr Gln Ala Val Met Val Val Ser Leu Pro Tyr Gln Asp Gln  
145 150 155 160

Val Ser Arg Leu Met Asp Arg Asp Gly Ile Asp Arg Asp Gln Ala Leu  
165 170 175

Arg Lys Ile Gln Ala Gln Met Ser Leu Glu Glu Lys Val Lys Leu Ala  
180 185 190



Asp Tyr Val Ile Asp Asn Ser Gly Ser Lys Glu Glu Ser Arg Gln Gln  
 195 200 205

Val Glu Ala Trp Leu Asp Gln Lys Gly Phe Lys Asn Leu  
 210 215 220

<210> 95  
 <211> 1335  
 <212> DNA  
 <213> *Alloiococcus otitidis*

<220>  
 <221> CDS  
 <222> (4)..(1335)  
 <223>

<400> 95  
 ggt atg gac caa gac acc atc tat cac ttt gtt ggc att aaa gga tct 48  
 Met Asp Gln Asp Thr Ile Tyr His Phe Val Gly Ile Lys Gly Ser  
 1 5 10 15

ggc atg agt tca ctt gcc act atc ttg ttt gac aag ggc tta aat gtc 96  
 Gly Met Ser Ser Leu Ala Thr Ile Leu Phe Asp Lys Gly Leu Asn Val  
 20 25 30

caa gga tct gat gtc aaa aag tat ttc ttt acc caa aaa agc tta gaa 144  
 Gln Gly Ser Asp Val Lys Lys Tyr Phe Phe Thr Gln Lys Ser Leu Glu  
 35 40 45

gaa aaa aat ata aac att tta gaa ttt gac cct gat aac atc aaa cca 192  
 Glu Lys Asn Ile Asn Ile Leu Glu Phe Asp Pro Asp Asn Ile Lys Pro  
 50 55 60

ggt atg acc ctg ata gca ggc aat gcc ttt gga gac aac cat ccc gag 240  
 Gly Met Thr Leu Ile Ala Gly Asn Ala Phe Gly Asp Asn His Pro Glu  
 65 70 75

ctg gtc cga ggt cga gag ctg ggt tta gaa atc atc cgc tac cat gat 288  
 Leu Val Arg Gly Arg Glu Leu Gly Leu Glu Ile Ile Arg Tyr His Asp  
 80 85 90 95

ttt atc ggt gac ctt atc gaa cac ttt act tcc atc gct att acc ggg 336  
 Phe Ile Gly Asp Leu Ile Glu His Phe Thr Ser Ile Ala Ile Thr Gly  
 100 105 110

tct cac ggt aag acc tcc aca act ggt ttg atg gcc cat gtt ttc tcc 384  
 Ser His Gly Lys Thr Ser Thr Thr Gly Leu Met Ala His Val Phe Ser  
 115 120 125

ggt att gat agc acc tcc tac tta att gga gat ggg acc ggc cat ggg 432  
 Gly Ile Asp Ser Thr Ser Tyr Leu Ile Gly Asp Gly Thr Gly His Gly  
 130 135 140

gaa aaa ggt gcc aag tat ttt gtc ttg gaa gcc tgc gaa tac aag cgg 480  
 Glu Lys Gly Ala Lys Tyr Phe Val Leu Glu Ala Cys Glu Tyr Lys Arg

145	150	155	
cac ttt ttg gcc tac cga ccg gac tat gcg gtt atg acc aat att gac			528
His Phe Leu Ala Tyr Arg Pro Asp Tyr Ala Val Met Thr Asn Ile Asp			
160	165	170	175
ttt gac cac ccg gac tat tac aag tct att gaa gat gtc caa gtg gcc			576
Phe Asp His Pro Asp Tyr Tyr Lys Ser Ile Glu Asp Val Gln Val Ala			
180	185		190
ttt gat gaa ttc agc cac cag gtc aaa aaa tac ctc ttt gcc tgc ggg			624
Phe Asp Glu Phe Ser His Gln Val Lys Lys Tyr Leu Phe Ala Cys Gly			
195	200		205
gac gac caa cgt ctt cgg cag gtc aaa gcc cag gtg ccg gtc att tac			672
Asp Asp Gln Arg Leu Arg Gln Val Lys Ala Gln Val Pro Val Ile Tyr			
210	215		220
tac ggt cta aat gaa gac aat gac ttt gtg gct aaa aac atc gac cga			720
Tyr Gly Leu Asn Glu Asp Asn Asp Phe Val Ala Lys Asn Ile Asp Arg			
225	230		235
agt cgt gaa ggg tct gcc ttc gac ctt tat att aag gga gaa ttt tac			768
Ser Arg Glu Gly Ser Ala Phe Asp Leu Tyr Ile Lys Gly Glu Phe Tyr			
240	245		250
aaa cac ttc acc atc cca acc tat ggc aac cac aat att caa aat gcc			816
Lys His Phe Thr Ile Pro Thr Tyr Gly Asn His Asn Ile Gln Asn Ala			
260	265		270
ttg gcg gtt ata gca gta gct tac tac gaa ggg tta gac caa gat ttg			864
Leu Ala Val Ile Ala Val Ala Tyr Tyr Glu Gly Leu Asp Gln Asp Leu			
275	280		285
gtt gcc caa aga ttg gct aat ttt gct ggg gtg aaa cgc cgg ttt acc			912
Val Ala Gln Arg Leu Ala Asn Phe Ala Gly Val Lys Arg Arg Phe Thr			
290	295		300
gag aag gtg gtc ggg gac act act att atc gat gac tat gct cac cac			960
Glu Lys Val Val Gly Asp Thr Thr Ile Ile Asp Asp Tyr Ala His His			
305	310		315
cct gct gaa ata agg gca acg att gat gcg gcc cgg caa aaa tac ccg			1008
Pro Ala Glu Ile Arg Ala Thr Ile Asp Ala Ala Arg Gln Lys Tyr Pro			
320	325		330
gac aag gac att gtg acg gtc ttc cag ccc cac acc ttt acc cgg aca			1056
Asp Lys Asp Ile Val Thr Val Phe Gln Pro His Thr Phe Thr Arg Thr			
340	345		350
gtc gcc ctc cta gat gaa ttt gcc cag gcc ttg gac ttg gca gac cag			1104
Val Ala Leu Leu Asp Glu Phe Ala Gln Ala Leu Asp Leu Ala Asp Gln			
355	360		365
gtt tac ttg tgt gat atc ttt aat tca gct aga gaa aag tca ggc gat			1152
Val Tyr Leu Cys Asp Ile Phe Asn Ser Ala Arg Glu Lys Ser Gly Asp			
370	375		380

att tcc atc caa gat ctt ttg gct aaa acc agc aag gcc gac cag gtg 1200  
 ile Ser ile Gln Asp Leu Leu Ala Lys Thr Ser Lys Ala Asp Gln Val  
 385 390 395

att gag gaa gac gat gtg tct cct ctg ctt gac caa cat ggg caa gtg 1248  
 ile Glu Glu Asp Asp Val Ser Pro Leu Leu Asp Gln His Gly Gln Val  
 400 405 410 415

att att ttc atg gga gca gga gac atc agc aag ttt gaa aaa gcc tat 1296  
 ile ile Phe Met Gly Ala Gly Asp ile Ser Lys Phe Glu Lys Ala Tyr  
 420 425 430

gaa agc ttg ttg agc tca acc tac cac tcc cag gtc taa 1335  
 Glu Ser Leu Leu Ser Ser Thr Tyr His Ser Gln Val  
 435 440

<210> 96  
 <211> 443  
 <212> PRT  
 <213> *Alloiococcus otitidis*

<400> 96  
 Met Asp Gln Asp Thr ile Tyr His Phe Val Gly ile Lys Gly Ser Gly  
 1 5 10 15

Met Ser Ser Leu Ala Thr ile Leu Phe Asp Lys Gly Leu Asn Val Gln  
 20 25 30

Gly Ser Asp Val Lys Lys Tyr Phe Phe Thr Gln Lys Ser Leu Glu Glu  
 35 40 45

Lys Asn ile Asn ile Leu Glu Phe Asp Pro Asp Asn ile Lys Pro Gly  
 50 55 60

Met Thr Leu ile Ala Gly Asn Ala Phe Gly Asp Asn His Pro Glu Leu  
 65 70 75 80

Val Arg Gly Arg Glu Leu Gly Leu Glu ile ile Arg Tyr His Asp Phe  
 85 90 95

ile Gly Asp Leu ile Glu His Phe Thr Ser ile Ala ile Thr Gly Ser  
 100 105 110

His Gly Lys Thr Ser Thr Thr Gly Leu Met Ala His Val Phe Ser Gly  
 115 120 125

ile Asp Ser Thr Ser Tyr Leu ile Gly Asp Gly Thr Gly His Gly Glu

130

135

140

Lys Gly Ala Lys Tyr Phe Val Leu Glu Ala Cys Glu Tyr Lys Arg His  
145 150 155 160

Phe Leu Ala Tyr Arg Pro Asp Tyr Ala Val Met Thr Asn Ile Asp Phe  
165 170 175

Asp His Pro Asp Tyr Tyr Lys Ser Ile Glu Asp Val Gln Val Ala Phe  
180 185 190

Asp Glu Phe Ser His Gln Val Lys Lys Tyr Leu Phe Ala Cys Gly Asp  
195 200 205

Asp Gln Arg Leu Arg Gln Val Lys Ala Gln Val Pro Val Ile Tyr Tyr  
210 215 220

Gly Leu Asn Glu Asp Asn Asp Phe Val Ala Lys Asn Ile Asp Arg Ser  
225 230 235 240

Arg Glu Gly Ser Ala Phe Asp Leu Tyr Ile Lys Gly Glu Phe Tyr Lys  
245 250 255

His Phe Thr Ile Pro Thr Tyr Gly Asn His Asn Ile Gln Asn Ala Leu  
260 265 270

Ala Val Ile Ala Val Ala Tyr Tyr Glu Gly Leu Asp Gln Asp Leu Val  
275 280 285

Ala Gln Arg Leu Ala Asn Phe Ala Gly Val Lys Arg Arg Phe Thr Glu  
290 295 300

Lys Val Val Gly Asp Thr Thr Ile Ile Asp Asp Tyr Ala His His Pro  
305 310 315 320

Ala Glu Ile Arg Ala Thr Ile Asp Ala Ala Arg Gln Lys Tyr Pro Asp  
325 330 335

Lys Asp Ile Val Thr Val Phe Gln Pro His Thr Phe Thr Arg Thr Val  
340 345 350

Ala Leu Leu Asp Glu Phe Ala Gln Ala Leu Asp Leu Ala Asp Gln Val  
355 360 365

Tyr Leu Cys Asp Ile Phe Asn Ser Ala Arg Glu Lys Ser Gly Asp Ile  
 370 375 380

Ser Ile Gln Asp Leu Leu Ala Lys Thr Ser Lys Ala Asp Gln Val Ile  
 385 390 395 400

Glu Glu Asp Asp Val Ser Pro Leu Leu Asp Gln His Gly Gln Val Ile  
 405 410 415

Ile Phe Met Gly Ala Gly Asp Ile Ser Lys Phe Glu Lys Ala Tyr Glu  
 420 425 430

Ser Leu Leu Ser Ser Thr Tyr His Ser Gln Val  
 435 440

<210> 97

<211> 1050

<212> DNA

<213> *Alloiococcus otitidis*

<220>

<221> CDS

<222> (19)..(1050)

<223>

<400> 97

acaaaattat ttacgtgt atg gag gaa tta ata gtg cca tta tta gac tta 51  
 Met Glu Glu Leu Ile Val Pro Leu Leu Asp Leu  
 1 5 10

aat gac cat gac cgc gtt cag gaa tat gag gac ttt gtc caa aac cac 99  
 Asn Asp His Asp Arg Val Gln Glu Tyr Glu Asp Phe Val Gln Asn His  
 15 20 25

ccc cag ggc cac ctg atg cag tct acc aaa tgg atc cag gtt aag gaa 147  
 Pro Gln Gly His Leu Met Gln Ser Thr Lys Trp Ile Gln Val Lys Glu  
 30 35 40

ggc tgg gac ggt gac tat gtt tac ctt acc gat gac caa gac cgg atc 195  
 Gly Trp Asp Gly Asp Tyr Val Tyr Leu Thr Asp Asp Gln Asp Arg Ile  
 45 50 55

aag gca tgc ttg tcc att cta tca gtc aaa aat gac gga gaa cat gcc 243  
 Lys Ala Cys Leu Ser Ile Leu Ser Val Lys Asn Asp Gly Glu His Ala  
 60 65 70 75

ttc tta tat gcg cca aga ggg ccg gtt tgt gac ttt cat gat aca gac 291  
 Phe Leu Tyr Ala Pro Arg Gly Pro Val Cys Asp Phe His Asp Thr Asp  
 80 85 90

ttg gtg acc gac tta att aag gaa gcc caa gtc gta gcg gac aag cac Leu Val Thr Asp Leu Ile Lys Glu Ala Gln Val Val Ala Asp Lys His 95 100 105	339
aag gcc ttt ttg ttg cgg atg gac ccg gaa acc ctt cat gat cct gac Lys Ala Phe Leu Leu Arg Met Asp Pro Glu Thr Leu His Asp Pro Asp 110 115 120	387
ctg gtc gaa aaa tac cgc gat tta ggc tat act ttc cgg tca gct gag Leu Val Glu Lys Tyr Arg Asp Leu Gly Tyr Thr Phe Arg Ser Ala Glu 125 130 135	435
caa gaa gat gaa cac gtc ttc tcc aac ccc cgc ttc cac atg atg acg Gln Glu Asp Glu His Val Phe Ser Asn Pro Arg Phe His Met Met Thr 140 145 150 155	483
gac tta agg ggt cat gat gaa gaa agc ttg ctg atg gcc ttc acc agc Asp Leu Arg Gly His Asp Glu Glu Ser Leu Leu Met Ala Phe Thr Ser 160 165 170	531
aat aac cgg cgc aag atc cgc aaa act tac aaa aat aac ctc cag acc Asn Asn Arg Arg Lys Ile Arg Lys Thr Tyr Lys Asn Asn Leu Gln Thr 175 180 185	579
cac tat ctg acc gtg gat gat gag ggt tat gac cag gcc ttg gat gac His Tyr Leu Thr Val Asp Asp Glu Gly Tyr Asp Gln Ala Leu Asp Asp 190 195 200	627
ttt tat gaa ttg acc caa ata atg gca gaa cgg caa ggg att act cac Phe Tyr Glu Leu Thr Gln Ile Met Ala Glu Arg Gln Gly Ile Thr His 205 210 215	675
cgg ccc aaa gac tac ttt gac cgg tta atg cac agc ttt gag gat gct Arg Pro Lys Asp Tyr Phe Asp Arg Leu Met His Ser Phe Glu Asp Ala 220 225 230 235	723
aaa ttg ttc cag acc tac cac gaa gat gac ctc cta gct act tgt atc Lys Leu Phe Gln Thr Tyr His Glu Asp Asp Leu Leu Ala Thr Cys Ile 240 245 250	771
ttg gtg agc tat aat aaa aaa tcc ttc tac atg tat gca gct tct tcc Leu Val Ser Tyr Asn Lys Lys Ser Phe Tyr Met Tyr Ala Ala Ser Ser 255 260 265	819
aac aaa aaa cga aat tta aat ggg tct ttg caa gaa aat tac gaa gcc Asn Lys Lys Arg Asn Leu Asn Gly Ser Leu Gln Glu Asn Tyr Glu Ala 270 275 280	867
atg aag tat gcc ttg gcc cga gga agc gaa gaa tat gat atg ggt ggg Met Lys Tyr Ala Leu Ala Arg Gly Ser Glu Glu Tyr Asp Met Gly Gly 285 290 295	915
gtc ttt ggc ttt gac aag tcg gac ggc ctc tac cgg ttt aaa aaa atc Val Phe Gly Phe Asp Lys Ser Asp Gly Leu Tyr Arg Phe Lys Lys Ile 300 305 310 315	963
ttt acc ggt cat gaa ggg ctg aaa gaa ttt atg ggt gaa ttg gat gtg	1011

Phe Thr Gly His Glu Gly Leu Lys Glu Phe Met Gly Glu Leu Asp Val  
 320 325 330

gtc tat gac caa gac cta tac gac gat ttt att tct taa 1050  
 Val Tyr Asp Gln Asp Leu Tyr Asp Asp Phe Ile Ser  
 335 340

<210> 98

<211> 343

<212> PRT

<213> *Alloiococcus otitidis*

<400> 98

Met Glu Glu Leu Ile Val Pro Leu Leu Asp Leu Asn Asp His Asp Arg  
 1 5 10 15

Val Gln Glu Tyr Glu Asp Phe Val Gln Asn His Pro Gln Gly His Leu  
 20 25 30

Met Gln Ser Thr Lys Trp Ile Gln Val Lys Glu Gly Trp Asp Gly Asp  
 35 40 45

Tyr Val Tyr Leu Thr Asp Asp Gln Asp Arg Ile Lys Ala Cys Leu Ser  
 50 55 60

Ile Leu Ser Val Lys Asn Asp Gly Glu His Ala Phe Leu Tyr Ala Pro  
 65 70 75 80

Arg Gly Pro Val Cys Asp Phe His Asp Thr Asp Leu Val Thr Asp Leu  
 85 90 95

Ile Lys Glu Ala Gln Val Val Ala Asp Lys His Lys Ala Phe Leu Leu  
 100 105 110

Arg Met Asp Pro Glu Thr Leu His Asp Pro Asp Leu Val Glu Lys Tyr  
 115 120 125

Arg Asp Leu Gly Tyr Thr Phe Arg Ser Ala Glu Gln Glu Asp Glu His  
 130 135 140

Val Phe Ser Asn Pro Arg Phe His Met Met Thr Asp Leu Arg Gly His  
 145 150 155 160

Asp Glu Glu Ser Leu Leu Met Ala Phe Thr Ser Asn Asn Arg Arg Lys  
 165 170 175

Ile Arg Lys Thr Tyr Lys Asn Asn Leu Gln Thr His Tyr Leu Thr Val  
180 185 190

Asp Asp Glu Gly Tyr Asp Gln Ala Leu Asp Asp Phe Tyr Glu Leu Thr  
195 200 205

Gln Ile Met Ala Glu Arg Gln Gly Ile Thr His Arg Pro Lys Asp Tyr  
210 215 220

Phe Asp Arg Leu Met His Ser Phe Glu Asp Ala Lys Leu Phe Gln Thr  
225 230 235 240

Tyr His Glu Asp Asp Leu Leu Ala Thr Cys Ile Leu Val Ser Tyr Asn  
245 250 255

Lys Lys Ser Phe Tyr Met Tyr Ala Ala Ser Ser Asn Lys Lys Arg Asn  
260 265 270

Leu Asn Gly Ser Leu Gln Glu Asn Tyr Glu Ala Met Lys Tyr Ala Leu  
275 280 285

Ala Arg Gly Ser Glu Glu Tyr Asp Met Gly Gly Val Phe Gly Phe Asp  
290 295 300

Lys Ser Asp Gly Leu Tyr Arg Phe Lys Lys Ile Phe Thr Gly His Glu  
305 310 315 320

Gly Leu Lys Glu Phe Met Gly Glu Leu Asp Val Val Tyr Asp Gln Asp  
325 330 335

Leu Tyr Asp Asp Phe Ile Ser  
340

<210> 99

<211> 2244

<212> DNA

<213> *Alloiococcus otitidis*

<220>

<221> CDS

<222> (22) .. (2244)

<223>

<400> 99

ttacgtgaaa ggaagacttg c atg ggc cta gca aaa gat att tta ggc aaa



atg	aat	gac	aaa	caa	aaa	caa	gcg	gtc	atg	acc	act	gat	ggc	cct	ctc	99
Met	Asn	Asp	Lys	Gln	Lys	Gln	Ala	Val	Met	Thr	Thr	Asp	Gly	Pro	Leu	
			15						20					25		
ttg	atc	atg	gct	ggg	gca	gga	tct	ggc	aag	acc	cgg	gtc	tta	acc	cac	147
Leu	Ile	Met	Ala	Gly	Ala	Gly	Ser	Gly	Lys	Thr	Arg	Val	Leu	Thr	His	
			30					35					40			
cgg	ata	gct	tac	ttg	atc	caa	gaa	aaa	ggg	gtt	aat	cct	tggt	aat	atc	195
Arg	Ile	Ala	Tyr	Leu	Ile	Gln	Glu	Lys	Gly	Val	Asn	Pro	Trp	Asn	Ile	
		45					50					55				
tta	gcc	atc	acc	ttt	acc	aac	aag	gcg	gct	ggc	gag	atg	aaa	gac	cggt	243
Leu	Ala	Ile	Thr	Phe	Thr	Asn	Lys	Ala	Ala	Gly	Glu	Met	Lys	Asp	Arg	
	60					65					70					
gtc	cag	aaa	ctg	gtt	agc	cag	gga	gga	tct	gga	gtt	tggt	gtc	tcgt	act	291
Val	Gln	Lys	Leu	Val	Ser	Gln	Gly	Gly	Ser	Gly	Val	Trp	Val	Ser	Thr	
75					80					85					90	
ttc	cac	tct	atg	tgt	gtt	cggt	att	cta	aga	agg	gac	gggt	gac	caa	att	339
Phe	His	Ser	Met	Cys	Val	Arg	Ile	Leu	Arg	Arg	Asp	Gly	Asp	Gln	Ile	
				95				100						105		
ggc	tat	aac	cgt	gcc	ttc	acc	att	gct	gac	cct	agt	gaa	cag	aaa	agt	387
Gly	Tyr	Asn	Arg	Ala	Phe	Thr	Ile	Ala	Asp	Pro	Ser	Glu	Gln	Lys	Ser	
			110					115					120			
ttg	atg	aag	cag	gtc	tta	aaa	gac	ttg	aat	att	gat	cct	aaa	cgt	tac	435
Leu	Met	Lys	Gln	Val	Leu	Lys	Asp	Leu	Asn	Ile	Asp	Pro	Lys	Arg	Tyr	
		125					130					135				
aac	ccc	aag	gcgt	ata	ttg	gcc	gag	att	tcc	aat	gcc	aaa	aat	gac	ctc	483
Asn	Pro	Lys	Ala	Ile	Leu	Ala	Glu	Ile	Ser	Asn	Ala	Lys	Asn	Asp	Leu	
	140					145					150					
ttg	gat	gag	caa	acc	tac	cggt	aaa	caa	gct	gat	gac	tat	ttt	aag	gaa	531
Leu	Asp	Glu	Gln	Thr	Tyr	Arg	Lys	Gln	Ala	Asp	Asp	Tyr	Phe	Lys	Glu	
155					160					165					170	
gtg	gtg	gct	gac	tgc	tac	gat	gct	tac	caa	aga	cag	ctc	cggt	cag	tct	579
Val	Val	Ala	Asp	Cys	Tyr	Asp	Ala	Tyr	Gln	Arg	Gln	Leu	Arg	Gln	Ser	
				175					180					185		
gag	gcc	atg	gac	ttt	gac	gac	ctgt	att	atg	caa	acc	gtc	cgt	ctc	ttc	627
Glu	Ala	Met	Asp	Phe	Asp	Asp	Leu	Ile	Met	Gln	Thr	Val	Arg	Leu	Phe	
			190					195					200			
aag	gaa	aag	ccc	gat	acc	ctgt	tct	tac	tac	cag	gcc	aag	ttc	cag	tat	675
Lys	Glu	Lys	Pro	Asp	Thr	Leu	Ser	Tyr	Tyr	Gln	Ala	Lys	Phe	Gln	Tyr	
		205														

220	225	230	
gtt caa ctg cta gcc caa cgc ttt aaa aat gtt tgc gtc gtg gga gat Val Gln Leu Leu Ala Gln Arg Phe Lys Asn Val Cys Val Val Gly Asp 235 240 245 250			771
gct gac cag tct att tat ggt tgg cgg ggg gct gat atg gga aat att Ala Asp Gln Ser Ile Tyr Gly Trp Arg Gly Ala Asp Met Gly Asn Ile 255 260 265			819
ttg aat ttc gaa aaa gac tat cca gaa gcc caa acc atc ttt ttg gaa Leu Asn Phe Glu Lys Asp Tyr Pro Glu Ala Gln Thr Ile Phe Leu Glu 270 275 280			867
caa aat tac cgg tca acc aag tct ata atc agg gca gcc aat gat gtt Gln Asn Tyr Arg Ser Thr Lys Ser Ile Ile Arg Ala Ala Asn Asp Val 285 290 295			915
atc caa aac aat atc aac cgc cgg gac aag aat ttg tgg act gcc aac Ile Gln Asn Asn Ile Asn Arg Arg Asp Lys Asn Leu Trp Thr Ala Asn 300 305 310			963
gat gag ggg gac aag gtc agc tta tac gct gcc cgg agc gag cag gat Asp Glu Gly Asp Lys Val Ser Leu Tyr Ala Ala Arg Ser Glu Gln Asp 315 320 325 330			1011
gaa gcc cag ttt atc gta ggg acc atc cat gac cta aca gaa ggc aaa Glu Ala Gln Phe Ile Val Gly Thr Ile His Asp Leu Thr Glu Gly Lys 335 340 345			1059
aag gct ggc tat ggg gac atc gcc atc ctc tac cgg acc aat gcc atg Lys Ala Gly Tyr Gly Asp Ile Ala Ile Leu Tyr Arg Thr Asn Ala Met 350 355 360			1107
tcc cgg gtt att gaa gaa acc ttt atc aag tcg aat atc ccc tac aag Ser Arg Val Ile Glu Glu Thr Phe Ile Lys Ser Asn Ile Pro Tyr Lys 365 370 375			1155
atc gtc ggc gga acc ggc ttt tac caa aga aaa gaa atc cgt gac ctg Ile Val Gly Gly Thr Gly Phe Tyr Gln Arg Lys Glu Ile Arg Asp Leu 380 385 390			1203
att gcc tac cta acc cta gtg gct aac cca gct gat gac ctg tcc ttt Ile Ala Tyr Leu Thr Leu Val Ala Asn Pro Ala Asp Asp Leu Ser Phe 395 400 405 410			1251
tca cgg atc gtt aat gag ccc aaa aga ggg att gga ccc ggc acc ctg Ser Arg Ile Val Asn Glu Pro Lys Arg Gly Ile Gly Pro Gly Thr Leu 415 420 425			1299
gac aag tta cgc cag gct ggc cag gag atg ggt tgg tcg ctt tac gaa Asp Lys Leu Arg Gln Ala Gly Gln Glu Met Gly Trp Ser Leu Tyr Glu 430 435 440			1347
aca gct ctc aat gcg gat gct acc aac ctg cct agt cgg gct gtc aac Thr Ala Leu Asn Ala Asp Ala Thr Asn Leu Pro Ser Arg Ala Val Asn 445 450 455			1395

aga cta tta gac ttc agt caa atg att gaa aat ttc agg aaa atg acg	1443
Arg Leu Leu Asp Phe Ser Gln Met Ile Glu Asn Phe Arg Lys Met Thr	
460 465 470	
gaa tac tta ccg att act gat ttg acc gaa aaa atc tta gag gat act	1491
Glu Tyr Leu Pro Ile Thr Asp Leu Thr Glu Lys Ile Leu Glu Asp Thr	
475 480 485 490	
ggc tac caa aaa gcc tta gaa aaa gac cgg act ctt gaa tct cag gca	1539
Gly Tyr Gln Lys Ala Leu Glu Lys Asp Arg Thr Leu Glu Ser Gln Ala	
495 500 505	
agg tta gag aac cta cag gaa ttt tac tcc gtc acc cag gaa ttt gac	1587
Arg Leu Glu Asn Leu Gln Glu Phe Tyr Ser Val Thr Gln Glu Phe Asp	
510 515 520	
cag caa gaa gac gac aac aag tca ctc tta gcc ttc tta act gac ctt	1635
Gln Gln Glu Asp Asp Asn Lys Ser Leu Leu Ala Phe Leu Thr Asp Leu	
525 530 535	
tcc tta ttg tca cca gct gat gat gtt gaa gag ggt cgg ggc cag gtc	1683
Ser Leu Leu Ser Pro Ala Asp Asp Val Glu Glu Gly Arg Gly Gln Val	
540 545 550	
acc atg atg acc ctc cat gca gcc aag ggg ttg gaa ttc ccc tat gtc	1731
Thr Met Met Thr Leu His Ala Ala Lys Gly Leu Glu Phe Pro Tyr Val	
555 560 565 570	
ttt atc gct ggt atg gaa gag gga atc ttc ccc ttg tcc cgg gcg gct	1779
Phe Ile Ala Gly Met Glu Glu Gly Ile Phe Pro Leu Ser Arg Ala Ala	
575 580 585	
gaa gac ccg gaa agc ttg gaa gaa gag cga cga ctg gcc tat gta ggg	1827
Glu Asp Pro Glu Ser Leu Glu Glu Glu Arg Arg Leu Ala Tyr Val Gly	
590 595 600	
att acc cgg gct gag cag gcc ctc tac cta acc cgt gcc atg atg cgc	1875
Ile Thr Arg Ala Glu Gln Ala Leu Tyr Leu Thr Arg Ala Met Met Arg	
605 610 615	
caa ctc tat ggc cgg acc cag gct aat ccc aaa tct cgc ttt tta tct	1923
Gln Leu Tyr Gly Arg Thr Gln Ala Asn Pro Lys Ser Arg Phe Leu Ser	
620 625 630	
gaa att tct tct gac ctg gtc caa gac ctt ggt gct aca act ggg tct	1971
Glu Ile Ser Ser Asp Leu Val Gln Asp Leu Gly Ala Thr Thr Gly Ser	
635 640 645 650	
ctt agc cag act ggg ggg aaa gtt agc cct aga cta gga ggc cgc aaa	2019
Leu Ser Gln Thr Gly Gly Lys Val Ser Pro Arg Leu Gly Gly Arg Lys	
655 660 665	
gcc agt ggt tat aag gct aat gct tgg tct cag caa tca gtt ggg gcg	2067
Ala Ser Gly Tyr Lys Ala Asn Ala Trp Ser Gln Gln Ser Val Gly Ala	
670 675 680	

act ggg gct gaa aaa gaa gac tgg gaa gtt ggt gac aag gtc cac cac 2115  
 Thr Gly Ala Glu Lys Glu Asp Trp Glu Val Gly Asp Lys Val His His  
 685 690 695

aaa aaa tgg ggc caa gga acc att att gag att aaa ggt tct ggc tcg 2163  
 Lys Lys Trp Gly Gln Gly Thr Ile Ile Glu Ile Lys Gly Ser Gly Ser  
 700 705 710

gac ctc cag ctc aac att gcc ttt cca gat gaa ggg atc aag ccc ttg 2211  
 Asp Leu Gln Leu Asn Ile Ala Phe Pro Asp Glu Gly Ile Lys Pro Leu  
 715 720 725 730

cta gcc agt ttt gcc ccc atc gaa aag att tag 2244  
 Leu Ala Ser Phe Ala Pro Ile Glu Lys Ile  
 735 740

<210> 100

<211> 740

<212> PRT

<213> *Alloiococcus otitidis*

<400> 100

Met Gly Leu Ala Lys Asp Ile Leu Gly Lys Met Asn Asp Lys Gln Lys  
 1 5 10 15

Gln Ala Val Met Thr Thr Asp Gly Pro Leu Leu Ile Met Ala Gly Ala  
 20 25 30

Gly Ser Gly Lys Thr Arg Val Leu Thr His Arg Ile Ala Tyr Leu Ile  
 35 40 45

Gln Glu Lys Gly Val Asn Pro Trp Asn Ile Leu Ala Ile Thr Phe Thr  
 50 55 60

Asn Lys Ala Ala Gly Glu Met Lys Asp Arg Val Gln Lys Leu Val Ser  
 65 70 75 80

Gln Gly Gly Ser Gly Val Trp Val Ser Thr Phe His Ser Met Cys Val  
 85 90 95

Arg Ile Leu Arg Arg Asp Gly Asp Gln Ile Gly Tyr Asn Arg Ala Phe  
 100 105 110

Thr Ile Ala Asp Pro Ser Glu Gln Lys Ser Leu Met Lys Gln Val Leu  
 115 120 125

Lys Asp Leu Asn Ile Asp Pro Lys Arg Tyr Asn Pro Lys Ala Ile Leu  
 130 135 140

Ala Glu Ile Ser Asn Ala Lys Asn Asp Leu Leu Asp Glu Gln Thr Tyr  
 145 150 155 160

Arg Lys Gln Ala Asp Asp Tyr Phe Lys Glu Val Val Ala Asp Cys Tyr  
 165 170 175

Asp Ala Tyr Gln Arg Gln Leu Arg Gln Ser Glu Ala Met Asp Phe Asp  
 180 185 190

Asp Leu Ile Met Gln Thr Val Arg Leu Phe Lys Glu Lys Pro Asp Thr  
 195 200 205

Leu Ser Tyr Tyr Gln Ala Lys Phe Gln Tyr Ile His Val Asp Glu Tyr  
 210 215 220

Gln Asp Thr Asn Gln Ala Gln Tyr Gln Leu Val Gln Leu Leu Ala Gln  
 225 230 235 240

Arg Phe Lys Asn Val Cys Val Val Gly Asp Ala Asp Gln Ser Ile Tyr  
 245 250 255

Gly Trp Arg Gly Ala Asp Met Gly Asn Ile Leu Asn Phe Glu Lys Asp  
 260 265 270

Tyr Pro Glu Ala Gln Thr Ile Phe Leu Glu Gln Asn Tyr Arg Ser Thr  
 275 280 285

Lys Ser Ile Ile Arg Ala Ala Asn Asp Val Ile Gln Asn Asn Ile Asn  
 290 295 300

Arg Arg Asp Lys Asn Leu Trp Thr Ala Asn Asp Glu Gly Asp Lys Val  
 305 310 315 320

Ser Leu Tyr Ala Ala Arg Ser Glu Gln Asp Glu Ala Gln Phe Ile Val  
 325 330 335

Gly Thr Ile His Asp Leu Thr Glu Gly Lys Lys Ala Gly Tyr Gly Asp  
 340 345 350

Ile Ala Ile Leu Tyr Arg Thr Asn Ala Met Ser Arg Val Ile Glu Glu  
 355 360 365

Thr Phe Ile Lys Ser Asn Ile Pro Tyr Lys Ile Val Gly Gly Thr Gly  
370 375 380

Phe Tyr Gln Arg Lys Glu Ile Arg Asp Leu Ile Ala Tyr Leu Thr Leu  
385 390 395 400

Val Ala Asn Pro Ala Asp Asp Leu Ser Phe Ser Arg Ile Val Asn Glu  
405 410 415

Pro Lys Arg Gly Ile Gly Pro Gly Thr Leu Asp Lys Leu Arg Gln Ala  
420 425 430

Gly Gln Glu Met Gly Trp Ser Leu Tyr Glu Thr Ala Leu Asn Ala Asp  
435 440 445

Ala Thr Asn Leu Pro Ser Arg Ala Val Asn Arg Leu Leu Asp Phe Ser  
450 455 460

Gln Met Ile Glu Asn Phe Arg Lys Met Thr Glu Tyr Leu Pro Ile Thr  
465 470 475 480

Asp Leu Thr Glu Lys Ile Leu Glu Asp Thr Gly Tyr Gln Lys Ala Leu  
485 490 495

Glu Lys Asp Arg Thr Leu Glu Ser Gln Ala Arg Leu Glu Asn Leu Gln  
500 505 510

Glu Phe Tyr Ser Val Thr Gln Glu Phe Asp Gln Gln Glu Asp Asp Asn  
515 520 525

Lys Ser Leu Leu Ala Phe Leu Thr Asp Leu Ser Leu Leu Ser Pro Ala  
530 535 540

Asp Asp Val Glu Glu Gly Arg Gly Gln Val Thr Met Met Thr Leu His  
545 550 555 560

Ala Ala Lys Gly Leu Glu Phe Pro Tyr Val Phe Ile Ala Gly Met Glu  
565 570 575

Glu Gly Ile Phe Pro Leu Ser Arg Ala Ala Glu Asp Pro Glu Ser Leu  
580 585 590

Glu Glu Glu Arg Arg Leu Ala Tyr Val Gly Ile Thr Arg Ala Glu Gln  
 595 600 605

Ala Leu Tyr Leu Thr Arg Ala Met Met Arg Gln Leu Tyr Gly Arg Thr  
 610 615 620

Gln Ala Asn Pro Lys Ser Arg Phe Leu Ser Glu Ile Ser Ser Asp Leu  
 625 630 635 640

Val Gln Asp Leu Gly Ala Thr Thr Gly Ser Leu Ser Gln Thr Gly Gly  
 645 650 655

Lys Val Ser Pro Arg Leu Gly Gly Arg Lys Ala Ser Gly Tyr Lys Ala  
 660 665 670

Asn Ala Trp Ser Gln Gln Ser Val Gly Ala Thr Gly Ala Glu Lys Glu  
 675 680 685

Asp Trp Glu Val Gly Asp Lys Val His His Lys Lys Trp Gly Gln Gly  
 690 695 700

Thr Ile Ile Glu Ile Lys Gly Ser Gly Ser Asp Leu Gln Leu Asn Ile  
 705 710 715 720

Ala Phe Pro Asp Glu Gly Ile Lys Pro Leu Leu Ala Ser Phe Ala Pro  
 725 730 735

Ile Glu Lys Ile  
 740

<210> 101

<211> 1314

<212> DNA

<213> *Alloiococcus otitidis*

<220>

<221> CDS

<222> (4)..(1314)

<223>

<400> 101

agt atg gac aca atc gtc att caa gga gga gac aat cga ctt gag ggt 48  
 Met Asp Thr Ile Val Ile Gln Gly Gly Asp Asn Arg Leu Glu Gly  
 1 5 10 15

aca gtc aag gta gaa ggg gct aag aat gct gcc ctt cct atc ctg gct 96  
 Thr Val Lys Val Glu Gly Ala Lys Asn Ala Ala Leu Pro Ile Leu Ala

20	25	30	
gcc agt ctt tta cca gaa gat ggg aaa agt cac ctg tcc aat gtc ccc Ala Ser Leu Leu Pro Glu Asp Gly Lys Ser His Leu Ser Asn Val Pro 35 40 45			144
tta cta tct gat att tac acg atg caa gaa gtt ttg cgt tac tta aac Leu Leu Ser Asp Ile Tyr Thr Met Gln Glu Val Leu Arg Tyr Leu Asn 50 55 60			192
gtt gac att gac ttc gat gaa gac cac aac gaa atc gtc ata gat gct Val Asp Ile Asp Phe Asp Glu Asp His Asn Glu Ile Val Ile Asp Ala 65 70 75			240
aca gga gac ctg aat tcc aat acc cct tat gaa ttt atg agc aag atg Thr Gly Asp Leu Asn Ser Asn Thr Pro Tyr Glu Phe Met Ser Lys Met 80 85 90 95			288
cgg gct tcc atc att gtc atg ggt ccc tta cta gcc cgt aat ggt tat Arg Ala Ser Ile Ile Val Met Gly Pro Leu Leu Ala Arg Asn Gly Tyr 100 105 110			336
gcc aaa gtc gct ctt cct ggt ggt tgc gcg att ggg act cgt cct att Ala Lys Val Ala Leu Pro Gly Gly Cys Ala Ile Gly Thr Arg Pro Ile 115 120 125			384
gac ttg cac tta aaa ggc ttc cgg gct atg ggg gtc gat gtg gaa gtc Asp Leu His Leu Lys Gly Phe Arg Ala Met Gly Val Asp Val Glu Val 130 135 140			432
gaa gga ggt tat gtg atc gcc aca gtt caa gat gaa ctg gat ggc gct Glu Gly Gly Tyr Val Ile Ala Thr Val Gln Asp Glu Leu Asp Gly Ala 145 150 155			480
gat att tac ctt gac ttc cca agt gtt gga gct aca caa aat att ttg Asp Ile Tyr Leu Asp Phe Pro Ser Val Gly Ala Thr Gln Asn Ile Leu 160 165 170 175			528
atg gct gcc acc cgg gca aaa ggg aca aca gtc atc gag aat gca gct Met Ala Ala Thr Arg Ala Lys Gly Thr Thr Val Ile Glu Asn Ala Ala 180 185 190			576
cga gaa cct gaa att gtt gac ctt gcc aac tat ttg aac aag atg ggt Arg Glu Pro Glu Ile Val Asp Leu Ala Asn Tyr Leu Asn Lys Met Gly 195 200 205			624
gcc cgt att tac ggg gcc gga acc aat acc atg aga att gaa ggg gta Ala Arg Ile Tyr Gly Ala Gly Thr Asn Thr Met Arg Ile Glu Gly Val 210 215 220			672
gac aag cta gaa gct tgt gac cac tcc att att gcc gac cgg ata gaa Asp Lys Leu Glu Ala Cys Asp His Ser Ile Ile Ala Asp Arg Ile Glu 225 230 235			720
agt ggc acc ttt atg gta gca gct ggt gtc acc caa ggg aat gtc ttg Ser Gly Thr Phe Met Val Ala Ala Gly Val Thr Gln Gly Asn Val Leu 240 245 250 255			768



att gaa gac tgt atc gtc gaa cac aac cgc ccc tta att tcc aag tta 816  
 Ile Glu Asp Cys Ile Val Glu His Asn Arg Pro Leu Ile Ser Lys Leu  
 260 265 270

agt gaa atg ggc gtt caa ttt gag gaa gaa aaa acc ggc ctt cga gtc 864  
 Ser Glu Met Gly Val Gln Phe Glu Glu Glu Lys Thr Gly Leu Arg Val  
 275 280 285

atg gga cca gag acc tta cag gca aca gat gtt aaa acc ctg cct tat 912  
 Met Gly Pro Glu Thr Leu Gln Ala Thr Asp Val Lys Thr Leu Pro Tyr  
 290 295 300

cct ggc ttc cca act gat atg cag tca ccg atg aca gtc gcc caa acc 960  
 Pro Gly Phe Pro Thr Asp Met Gln Ser Pro Met Thr Val Ala Gln Thr  
 305 310 315

cta gct gag gga aga agc atc atg aga gaa acg gtc ttc gaa aac cgc 1008  
 Leu Ala Glu Gly Arg Ser Ile Met Arg Glu Thr Val Phe Glu Asn Arg  
 320 325 330 335

ttc atg cac atg gaa gag ctt cgt aaa atg gat gca caa ttt act gtc 1056  
 Phe Met His Met Glu Glu Leu Arg Lys Met Asp Ala Gln Phe Thr Val  
 340 345 350

gat ggc cag tcc ctt att atc gag ggg ggc aaa aaa ctc caa ggt gct 1104  
 Asp Gly Gln Ser Leu Ile Ile Glu Gly Gly Lys Lys Leu Gln Gly Ala  
 355 360 365

aga gtc cag tcc agt gac ttg cgg gct tca gct tcc ttg att att gct 1152  
 Arg Val Gln Ser Ser Asp Leu Arg Ala Ser Ala Ser Leu Ile Ile Ala  
 370 375 380

ggt tta gta gct gat ggt gtc acc aaa gta acc aat ctt aac cac tta 1200  
 Gly Leu Val Ala Asp Gly Val Thr Lys Val Thr Asn Leu Asn His Leu  
 385 390 395

gac cgg ggc tac tat aaa ttt cac gaa aaa tta cag caa tta ggt gct 1248  
 Asp Arg Gly Tyr Tyr Lys Phe His Glu Lys Leu Gln Gln Leu Gly Ala  
 400 405 410 415

tcc att gaa cga atc gac gag gaa att caa gtt gac cag gaa gcc agc 1296  
 Ser Ile Glu Arg Ile Asp Glu Glu Ile Gln Val Asp Gln Glu Ala Ser  
 420 425 430

ctc aaa aaa ggc gaa taa 1314  
 Leu Lys Lys Gly Glu  
 435

&lt;210&gt; 102

&lt;211&gt; 436

&lt;212&gt; PRT

&lt;213&gt; Alloiococcus otitidis

&lt;400&gt; 102

Met Asp Thr Ile Val Ile Gln Gly Gly Asp Asn Arg Leu Glu Gly Thr

**PCT/US02/36122**

1	5	10	15
Val Lys Val Glu Gly Ala Lys Asn Ala Ala Leu Pro Ile Leu Ala Ala	20	25	30
Ser Leu Leu Pro Glu Asp Gly Lys Ser His Leu Ser Asn Val Pro Leu	35	40	45
Leu Ser Asp Ile Tyr Thr Met Gln Glu Val Leu Arg Tyr Leu Asn Val	50	55	60
Asp Ile Asp Phe Asp Glu Asp His Asn Glu Ile Val Ile Asp Ala Thr	65	70	75
Gly Asp Leu Asn Ser Asn Thr Pro Tyr Glu Phe Met Ser Lys Met Arg	85	90	95
Ala Ser Ile Ile Val Met Gly Pro Leu Leu Ala Arg Asn Gly Tyr Ala	100	105	110
Lys Val Ala Leu Pro Gly Gly Cys Ala Ile Gly Thr Arg Pro Ile Asp	115	120	125
Leu His Leu Lys Gly Phe Arg Ala Met Gly Val Asp Val Glu Val Glu	130	135	140
Gly Gly Tyr Val Ile Ala Thr Val Gln Asp Glu Leu Asp Gly Ala Asp	145	150	155
Ile Tyr Leu Asp Phe Pro Ser Val Gly Ala Thr Gln Asn Ile Leu Met	165	170	175
Ala Ala Thr Arg Ala Lys Gly Thr Thr Val Ile Glu Asn Ala Ala Arg	180	185	190
Glu Pro Glu Ile Val Asp Leu Ala Asn Tyr Leu Asn Lys Met Gly Ala	195	200	205
Arg Ile Tyr Gly Ala Gly Thr Asn Thr Met Arg Ile Glu Gly Val Asp	210	215	220
Lys Leu Glu Ala Cys Asp His Ser Ile Ile Ala Asp Arg Ile Glu Ser	225	230	235
			240

Gly Thr Phe Met Val Ala Ala Gly Val Thr Gln Gly Asn Val Leu Ile  
 245 250 255

Glu Asp Cys Ile Val Glu His Asn Arg Pro Leu Ile Ser Lys Leu Ser  
 260 265 270

Glu Met Gly Val Gln Phe Glu Glu Glu Lys Thr Gly Leu Arg Val Met  
 275 280 285

Gly Pro Glu Thr Leu Gln Ala Thr Asp Val Lys Thr Leu Pro Tyr Pro  
 290 295 300

Gly Phe Pro Thr Asp Met Gln Ser Pro Met Thr Val Ala Gln Thr Leu  
 305 310 315 320

Ala Glu Gly Arg Ser Ile Met Arg Glu Thr Val Phe Glu Asn Arg Phe  
 325 330 335

Met His Met Glu Glu Leu Arg Lys Met Asp Ala Gln Phe Thr Val Asp  
 340 345 350

Gly Gln Ser Leu Ile Ile Glu Gly Gly Lys Lys Leu Gln Gly Ala Arg  
 355 360 365

Val Gln Ser Ser Asp Leu Arg Ala Ser Ala Ser Leu Ile Ile Ala Gly  
 370 375 380

Leu Val Ala Asp Gly Val Thr Lys Val Thr Asn Leu Asn His Leu Asp  
 385 390 395 400

Arg Gly Tyr Tyr Lys Phe His Glu Lys Leu Gln Gln Leu Gly Ala Ser  
 405 410 415

Ile Glu Arg Ile Asp Glu Glu Ile Gln Val Asp Gln Glu Ala Ser Leu  
 420 425 430

Lys Lys Gly Glu  
 435

<210> 103  
 <211> 1026  
 <212> DNA

&lt;213&gt; Alloioococcus otitidis

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (34)..(1026)

&lt;223&gt;

&lt;400&gt; 103

```

acagtttttaa tccagtttagc cacaaggtgg gat atg atg gac tta gca gaa aaa      54
                               Met Met Asp Leu Ala Glu Lys
                               1           5

caa gca ggg gtc tac caa ctt ttt gac cga atc ctg gcc aac cat gcc      102
Gln Ala Gly Val Tyr Gln Leu Phe Asp Arg Ile Leu Ala Asn His Ala
      10           15           20

ctc aag cat gcc tat ctt ttt gaa ggt ttg gcc gga tca ggc aaa ctg      150
Leu Lys His Ala Tyr Leu Phe Glu Gly Leu Ala Gly Ser Gly Lys Leu
      25           30           35

gag atg agc cgg tat att gcc aag aga ctg ttt tgc ccc aac caa gac      198
Glu Met Ser Arg Tyr Ile Ala Lys Arg Leu Phe Cys Pro Asn Gln Asp
      40           45           50           55

cag gga caa gct tgc caa gtt tgt ccc act tgc ttg cgc att gac cag      246
Gln Gly Gln Ala Cys Gln Val Cys Pro Thr Cys Leu Arg Ile Asp Gln
      60           65           70

ggg caa cac cct gat gtg gta gaa ata gcc cct gag ggg aag gga cgg      294
Gly Gln His Pro Asp Val Val Glu Ile Ala Pro Glu Gly Lys Gly Arg
      75           80           85

tcg att agg gta gac cgg gta cga cag gtc aag gat gcc cta agc aag      342
Ser Ile Arg Val Asp Arg Val Arg Gln Val Lys Asp Ala Leu Ser Lys
      90           95          100

tct ggt gtg gag agt caa aag aaa atg att atc ctt aac cag gct gat      390
Ser Gly Val Glu Ser Gln Lys Lys Met Ile Ile Leu Asn Gln Ala Asp
      105          110          115

aaa atg acc ccc agt gca gcc aac agc ctg ctt aaa ttt ctg gaa gag      438
Lys Met Thr Pro Ser Ala Ala Asn Ser Leu Leu Lys Phe Leu Glu Glu
      120          125          130          135

ccg gca ggg gat gtg act att ttc ttg tta gtt act agc cgg caa aac      486
Pro Ala Gly Asp Val Thr Ile Phe Leu Leu Val Thr Ser Arg Gln Asn
      140          145          150

ctt ttg cca act att gtt tcc cgc tgc cag gtt atc cag ttt gcc aag      534
Leu Leu Pro Thr Ile Val Ser Arg Cys Gln Val Ile Gln Phe Ala Lys
      155          160          165

cag gat tta aag act cgg att gag gac tta gtg gaa gcc ggt ttg tcc      582
Gln Asp Leu Lys Thr Arg Ile Glu Asp Leu Val Glu Ala Gly Leu Ser
      170          175          180

cag gaa gaa gcc cac ttg gcc agc cac ctc agc caa gac tta gac ttg      630

```

Gln Glu Glu Ala His Leu Ala Ser His Leu Ser Gln Asp Leu Asp Leu  
 185 190 195

gct aag tcc ctc att gag gaa gag gac ttg ctg gca gtc agt caa aaa 678  
 Ala Lys Ser Leu Ile Glu Glu Glu Asp Leu Leu Ala Val Ser Gln Lys  
 200 205 210 215

att tgg cag tgg ttt agc tat ctc atg aac caa gat gac ttg gcc ttt 726  
 Ile Trp Gln Trp Phe Ser Tyr Leu Met Asn Gln Asp Asp Leu Ala Phe  
 220 225 230

atc cta gtc caa aga gac tta atg gcc ttt atc caa gac cgg gat gac 774  
 Ile Leu Val Gln Arg Asp Leu Met Ala Phe Ile Gln Asp Arg Asp Asp  
 235 240 245

tgc cag atg gtt tgt gac tta atc ctc tac ctc ttc caa gac ctg ctc 822  
 Cys Gln Met Val Cys Asp Leu Ile Leu Tyr Leu Phe Gln Asp Leu Leu  
 250 255 260

cac tta cac tac cat tta gat agt ccg gcc tgc ttc gca ggc cac gaa 870  
 His Leu His Tyr His Leu Asp Ser Pro Ala Cys Phe Ala Gly His Glu  
 265 270 275

agt gac ctc cgc tac ttt atg gac ctg ctt tgc atc aag caa gtg tct 918  
 Ser Asp Leu Arg Tyr Phe Met Asp Leu Leu Ser Ile Lys Gln Val Ser  
 280 285 290 295

tat gcc atg caa gcc acc ctg caa gct aaa aga gaa gtg gac cac aat 966  
 Tyr Ala Met Gln Ala Thr Leu Gln Ala Lys Arg Glu Val Asp His Asn  
 300 305 310

gtg gcc agt cag gct gtt tta gaa ggc ttg act ttg gac ttg cag gaa 1014  
 Val Ala Ser Gln Ala Val Leu Glu Gly Leu Thr Leu Asp Leu Gln Glu  
 315 320 325

agt ata ggc taa 1026  
 Ser Ile Gly  
 330

&lt;210&gt; 104

&lt;211&gt; 330

&lt;212&gt; PRT

<213> *Alloiococcus otitidis*

&lt;400&gt; 104

Met Met Asp Leu Ala Glu Lys Gln Ala Gly Val Tyr Gln Leu Phe Asp  
 1 5 10 15

Arg Ile Leu Ala Asn His Ala Leu Lys His Ala Tyr Leu Phe Glu Gly  
 20 25 30

Leu Ala Gly Ser Gly Lys Leu Glu Met Ser Arg Tyr Ile Ala Lys Arg  
 35 40 45

Leu Phe Cys Pro Asn Gln Asp Gln Gly Gln Ala Cys Gln Val Cys Pro  
50 55 60

Thr Cys Leu Arg Ile Asp Gln Gly Gln His Pro Asp Val Val Glu Ile  
65 70 75 80

Ala Pro Glu Gly Lys Gly Arg Ser Ile Arg Val Asp Arg Val Arg Gln  
85 90 95

Val Lys Asp Ala Leu Ser Lys Ser Gly Val Glu Ser Gln Lys Lys Met  
100 105 110

Ile Ile Leu Asn Gln Ala Asp Lys Met Thr Pro Ser Ala Ala Asn Ser  
115 120 125

Leu Leu Lys Phe Leu Glu Glu Pro Ala Gly Asp Val Thr Ile Phe Leu  
130 135 140

Leu Val Thr Ser Arg Gln Asn Leu Leu Pro Thr Ile Val Ser Arg Cys  
145 150 155 160

Gln Val Ile Gln Phe Ala Lys Gln Asp Leu Lys Thr Arg Ile Glu Asp  
165 170 175

Leu Val Glu Ala Gly Leu Ser Gln Glu Glu Ala His Leu Ala Ser His  
180 185 190

Leu Ser Gln Asp Leu Asp Leu Ala Lys Ser Leu Ile Glu Glu Glu Asp  
195 200 205

Leu Leu Ala Val Ser Gln Lys Ile Trp Gln Trp Phe Ser Tyr Leu Met  
210 215 220

Asn Gln Asp Asp Leu Ala Phe Ile Leu Val Gln Arg Asp Leu Met Ala  
225 230 235 240

Phe Ile Gln Asp Arg Asp Asp Cys Gln Met Val Cys Asp Leu Ile Leu  
245 250 255

Tyr Leu Phe Gln Asp Leu Leu His Leu His Tyr His Leu Asp Ser Pro  
260 265 270

Ala Cys Phe Ala Gly His Glu Ser Asp Leu Arg Tyr Phe Met Asp Leu  
 275 280 285

Leu Ser Ile Lys Gln Val Ser Tyr Ala Met Gln Ala Thr Leu Gln Ala  
 290 295 300

Lys Arg Glu Val Asp His Asn Val Ala Ser Gln Ala Val Leu Glu Gly  
 305 310 315 320

Leu Thr Leu Asp Leu Gln Glu Ser Ile Gly  
 325 330

<210> 105

<211> 1785

<212> DNA

<213> *Alloioococcus otitidis*

<220>

<221> CDS

<222> (13)..(1785)

<223>

<400> 105

gaggggagag ct atg acc cac cag gcc tta tac cgg gta tgg cga ccg caa 51  
 Met Thr His Gln Ala Leu Tyr Arg Val Trp Arg Pro Gln  
 1 5 10

agt ttt gct gat gta tcc ggc cag cat gtg gtc acc aag acc cta aag 99  
 Ser Phe Ala Asp Val Ser Gly Gln His Val Val Thr Lys Thr Leu Lys  
 15 20 25

aat gcc att aaa aat gat aat acc agt cat gcc tac ctg ttt act gga 147  
 Asn Ala Ile Lys Asn Asp Asn Thr Ser His Ala Tyr Leu Phe Thr Gly  
 30 35 40 45

ccc cgg ggg acg ggc aag acc agt gtg gca aaa ata ttt gcc aag gcc 195  
 Pro Arg Gly Thr Gly Lys Thr Ser Val Ala Lys Ile Phe Ala Lys Ala  
 50 55 60

att aat tgc ccc tac tcg gat gat ggg gag cct tgt aat gaa tgt cag 243  
 Ile Asn Cys Pro Tyr Ser Asp Asp Gly Glu Pro Cys Asn Glu Cys Gln  
 65 70 75

att tgc cag gag atc acc cag ggt agt cta ggc gat gtc atc gaa atc 291  
 Ile Cys Gln Glu Ile Thr Gln Gly Ser Leu Gly Asp Val Ile Glu Ile  
 80 85 90

gat gcg gcc agc aat aat ggg gtg gaa gag att cgc gat att agg gaa 339  
 Asp Ala Ala Ser Asn Asn Gly Val Glu Glu Ile Arg Asp Ile Arg Glu  
 95 100 105

aag gct aat tat gcc cca act tcg gcc gtt tac aag gtc tac att atc 387  
 Lys Ala Asn Tyr Ala Pro Thr Ser Ala Val Tyr Lys Val Tyr Ile Ile

110	115	120	125	
gat gag gtc cat atg tta tcc tct ggg gcc ttt aac gcc ctc ttg aaa				435
Asp Glu Val His Met Leu Ser Ser Gly Ala Phe Asn Ala Leu Leu Lys				
130		135	140	
aca ctg gaa gag cct cca gcc aat gtg gtc ttt atc tta gca acg act				483
Thr Leu Glu Glu Pro Pro Ala Asn Val Val Phe Ile Leu Ala Thr Thr				
145		150	155	
gaa ccc cac aag att ccg gct acc att atc tcc cgg acc cag cgt ttt				531
Glu Pro His Lys Ile Pro Ala Thr Ile Ile Ser Arg Thr Gln Arg Phe				
160		165	170	
gat ttt aag cgg att gac aac cag gac atc atc gac cgc ttg att tat				579
Asp Phe Lys Arg Ile Asp Asn Gln Asp Ile Ile Asp Arg Leu Ile Tyr				
175		180	185	
atc tta gaa gaa gac cag gtc ccc tac agc aaa gaa gcc gtc cta agc				627
Ile Leu Glu Glu Asp Gln Val Pro Tyr Ser Lys Glu Ala Val Leu Ser				
190		195	200	205
cta gcc aat gca gcg gaa ggt ggg atg cgg gat gcc ttg agt atg ttg				675
Leu Ala Asn Ala Ala Glu Gly Gly Met Arg Asp Ala Leu Ser Met Leu				
210		215	220	
gac cag gcc tta agc ttt atg aca gat gag tta aca gaa gaa gtt gcc				723
Asp Gln Ala Leu Ser Phe Met Thr Asp Glu Leu Thr Glu Glu Val Ala				
225		230	235	
ctc cag att aca ggg agc att acc cag tct ctc ttg ctt gaa tac ttg				771
Leu Gln Ile Thr Gly Ser Ile Thr Gln Ser Leu Leu Leu Glu Tyr Leu				
240		245	250	
cag gtg att agc caa ggt cag acg gaa gaa gga ctc aag ctc ttg caa				819
Gln Val Ile Ser Gln Gly Gln Thr Glu Glu Gly Leu Lys Leu Leu Gln				
255		260	265	
gaa gtt tta ggg gaa ggc aag gac cct agc cgg ttt gtg gaa gac gct				867
Glu Val Leu Gly Glu Gly Lys Asp Pro Ser Arg Phe Val Glu Asp Ala				
270		275	280	285
att atg atg acc cgg gac ctc ttg ctt tac caa act agc caa ggc gat				915
Ile Met Met Thr Arg Asp Leu Leu Leu Tyr Gln Thr Ser Gln Gly Asp				
290		295	300	
aat ttt gtt cct aaa ttg gct cgc tta gac gac cag ttt gaa gac ctg				963
Asn Phe Val Pro Lys Leu Ala Arg Leu Asp Asp Gln Phe Glu Asp Leu				
305		310	315	
gcg aag gac ttg gac aag gag atg gcc tac cat att att gat gtc tta				1011
Ala Lys Asp Leu Asp Lys Glu Met Ala Tyr His Ile Ile Asp Val Leu				
320		325	330	
aac caa acc caa gac gat ctc cgc cta agc aac cat ggg gaa gtc tat				1059
Asn Gln Thr Gln Asp Asp Leu Arg Leu Ser Asn His Gly Glu Val Tyr				
335		340	345	



ttg gaa ata gcc acg gtc aag ctt agc cag cct tct tca gcc gtt cag Leu Glu Ile Ala Thr Val Lys Leu Ser Gln Pro Ser Ser Ala Val Gln 350 355 360 365	1107
acc atc cag gcc agc caa gtc aac atg gtg gac cag gat aat aaa gaa Thr Ile Gln Ala Ser Gln Val Asn Met Val Asp Gln Asp Asn Lys Glu 370 375 380	1155
gag att gcc caa ctg caa aac cag gtc aag tcc ctc cag caa agt att Glu Ile Ala Gln Leu Gln Asn Gln Val Lys Ser Leu Gln Gln Ser Ile 385 390 395	1203
caa aac ttg caa gct gga gcc aaa caa ggg cct aag caa aga gct aag Gln Asn Leu Gln Ala Gly Ala Lys Gln Gly Pro Lys Gln Arg Ala Lys 400 405 410	1251
tca aaa gct ggc ccc aag caa tct ggc cct ggc aag tct aga agc cac Ser Lys Ala Gly Pro Lys Gln Ser Gly Pro Gly Lys Ser Arg Ser His 415 420 425	1299
cgt cac cag caa ggc ttc aag gtt aac cgg aaa gcc gtt tac tct atc Arg His Gln Gln Gly Phe Lys Val Asn Arg Lys Ala Val Tyr Ser Ile 430 435 440 445	1347
ttg gac cag gcg acc cgt aaa gac ctg gac gac ctc caa gac ctc tgg Leu Asp Gln Ala Thr Arg Lys Asp Leu Asp Asp Leu Gln Asp Leu Trp 450 455 460	1395
cca gac ttg atc aat gtc ttg acc atc agt caa aag gct atc tta aac Pro Asp Leu Ile Asn Val Leu Thr Ile Ser Gln Lys Ala Ile Leu Asn 465 470 475	1443
aat tcc aaa cca gtt gct gct agt cca gag ggt ttg gtg gtg acc ttt Asn Ser Lys Pro Val Ala Ala Ser Pro Glu Gly Leu Val Val Thr Phe 480 485 490	1491
gaa tat gat att cta tgt gag aga gca gag tct gac gag acc ttg caa Glu Tyr Asp Ile Leu Cys Glu Arg Ala Glu Ser Asp Glu Thr Leu Gln 495 500 505	1539
acg gct atc ggc aat tac atc gaa aaa att atc ggc cgc cgt cca aga Thr Ala Ile Gly Asn Tyr Ile Glu Lys Ile Ile Gly Arg Arg Pro Arg 510 515 520 525	1587
ctg gtc tgt gtg cct gaa gac aag tgg ccg act atc cgc cgc gat ttt Leu Val Cys Val Pro Glu Asp Lys Trp Pro Thr Ile Arg Arg Asp Phe 530 535 540	1635
atc aag cag atg aaa aaa gaa gat ggc agt act aaa gct ggc caa gca Ile Lys Gln Met Lys Lys Glu Asp Gly Ser Thr Lys Ala Gly Gln Ala 545 550 555	1683
agt gac ggc aag tcg gat gat gac cca ggt caa gaa gac aac cag gcc Ser Asp Gly Lys Ser Asp Asp Asp Pro Gly Gln Glu Asp Asn Gln Ala 560 565 570	1731

ctt aac aag gct gtg gag ctt ttc ggt aaa gac aat att aca atc aaa 1779  
 Leu Asn Lys Ala Val Glu Leu Phe Gly Lys Asp Asn Ile Thr Ile Lys  
 575 580 585

gat taa 1785  
 Asp  
 590

<210> 106  
 <211> 590  
 <212> PRT  
 <213> Alloiococcus otitidis

<400> 106  
 Met Thr His Gln Ala Leu Tyr Arg Val Trp Arg Pro Gln Ser Phe Ala  
 1 5 10 15

Asp Val Ser Gly Gln His Val Val Thr Lys Thr Leu Lys Asn Ala Ile  
 20 25 30

Lys Asn Asp Asn Thr Ser His Ala Tyr Leu Phe Thr Gly Pro Arg Gly  
 35 40 45

Thr Gly Lys Thr Ser Val Ala Lys Ile Phe Ala Lys Ala Ile Asn Cys  
 50 55 60

Pro Tyr Ser Asp Asp Gly Glu Pro Cys Asn Glu Cys Gln Ile Cys Gln  
 65 70 75 80

Glu Ile Thr Gln Gly Ser Leu Gly Asp Val Ile Glu Ile Asp Ala Ala  
 85 90 95

Ser Asn Asn Gly Val Glu Glu Ile Arg Asp Ile Arg Glu Lys Ala Asn  
 100 105 110

Tyr Ala Pro Thr Ser Ala Val Tyr Lys Val Tyr Ile Ile Asp Glu Val  
 115 120 125

His Met Leu Ser Ser Gly Ala Phe Asn Ala Leu Leu Lys Thr Leu Glu  
 130 135 140

Glu Pro Pro Ala Asn Val Val Phe Ile Leu Ala Thr Thr Glu Pro His  
 145 150 155 160

Lys Ile Pro Ala Thr Ile Ile Ser Arg Thr Gln Arg Phe Asp Phe Lys  
 165 170 175

Arg Ile Asp Asn Gln Asp Ile Ile Asp Arg Leu Ile Tyr Ile Leu Glu  
180 185 190

Glu Asp Gln Val Pro Tyr Ser Lys Glu Ala Val Leu Ser Leu Ala Asn  
195 200 205

Ala Ala Glu Gly Gly Met Arg Asp Ala Leu Ser Met Leu Asp Gln Ala  
210 215 220

Leu Ser Phe Met Thr Asp Glu Leu Thr Glu Glu Val Ala Leu Gln Ile  
225 230 235 240

Thr Gly Ser Ile Thr Gln Ser Leu Leu Leu Glu Tyr Leu Gln Val Ile  
245 250 255

Ser Gln Gly Gln Thr Glu Glu Gly Leu Lys Leu Leu Gln Glu Val Leu  
260 265 270

Gly Glu Gly Lys Asp Pro Ser Arg Phe Val Glu Asp Ala Ile Met Met  
275 280 285

Thr Arg Asp Leu Leu Leu Tyr Gln Thr Ser Gln Gly Asp Asn Phe Val  
290 295 300

Pro Lys Leu Ala Arg Leu Asp Asp Gln Phe Glu Asp Leu Ala Lys Asp  
305 310 315 320

Leu Asp Lys Glu Met Ala Tyr His Ile Ile Asp Val Leu Asn Gln Thr  
325 330 335

Gln Asp Asp Leu Arg Leu Ser Asn His Gly Glu Val Tyr Leu Glu Ile  
340 345 350

Ala Thr Val Lys Leu Ser Gln Pro Ser Ser Ala Val Gln Thr Ile Gln  
355 360 365

Ala Ser Gln Val Asn Met Val Asp Gln Asp Asn Lys Glu Glu Ile Ala  
370 375 380

Gln Leu Gln Asn Gln Val Lys Ser Leu Gln Gln Ser Ile Gln Asn Leu  
385 390 395 400

Gln Ala Gly Ala Lys Gln Gly Pro Lys Gln Arg Ala Lys Ser Lys Ala  
405 410 415

Gly Pro Lys Gln Ser Gly Pro Gly Lys Ser Arg Ser His Arg His Gln  
420 425 430

Gln Gly Phe Lys Val Asn Arg Lys Ala Val Tyr Ser Ile Leu Asp Gln  
435 440 445

Ala Thr Arg Lys Asp Leu Asp Asp Leu Gln Asp Leu Trp Pro Asp Leu  
450 455 460

Ile Asn Val Leu Thr Ile Ser Gln Lys Ala Ile Leu Asn Asn Ser Lys  
465 470 475 480

Pro Val Ala Ala Ser Pro Glu Gly Leu Val Val Thr Phe Glu Tyr Asp  
485 490 495

Ile Leu Cys Glu Arg Ala Glu Ser Asp Glu Thr Leu Gln Thr Ala Ile  
500 505 510

Gly Asn Tyr Ile Glu Lys Ile Ile Gly Arg Arg Pro Arg Leu Val Cys  
515 520 525

Val Pro Glu Asp Lys Trp Pro Thr Ile Arg Arg Asp Phe Ile Lys Gln  
530 535 540

Met Lys Lys Glu Asp Gly Ser Thr Lys Ala Gly Gln Ala Ser Asp Gly  
545 550 555 560

Lys Ser Asp Asp Asp Pro Gly Gln Glu Asp Asn Gln Ala Leu Asn Lys  
565 570 575

Ala Val Glu Leu Phe Gly Lys Asp Asn Ile Thr Ile Lys Asp  
580 585 590